



SEQUENCE LISTING

(1) GENERAL INFORMATION:

- (i) APPLICANT: Christian, P. D., Gordon, K. H.J., Hanzlik, T. N.
- (ii) TITLE OF INVENTION: Insect Viruses and Their Uses in Protecting Plants
- (iii) NUMBER OF SEQUENCES: 57
- (iv) CORRESPONDENCE ADDRESS:
 - (A) ADDRESSEE: Flehr Hohbach Test Albritton & Herbert LLP
 - (B) STREET: Four Embarcadero Center, Suite 3400
 - (C) CITY: San Francisco
 - (D) STATE: California
 - (E) COUNTRY: United States
 - (F) ZIP: 94111-4187
- (v) COMPUTER READABLE FORM:
 - (A) MEDIUM TYPE: Floppy disk
 - (B) COMPUTER: IBM PC compatible
 - (C) OPERATING SYSTEM: PC-DOS/MS-DOS
 - (D) SOFTWARE: PatentIn Release #1.0, Version #1.30
- (vi) CURRENT APPLICATION DATA:
 - (A) APPLICATION NUMBER: US 09/234,238
 - (B) FILING DATE: 20-JAN-1999
 - (C) CLASSIFICATION:
- (vii) PRIOR APPLICATION DATA:
 - (A) APPLICATION NUMBER: US 08/485,355
 - (B) FILING DATE: 07-JUN-1995
- (vii) PRIOR APPLICATION DATA:
 - (A) APPLICATION NUMBER: US 08/440,522
 - (B) FILING DATE: 12-MAY-1995
- (vii) PRIOR APPLICATION DATA:
 - (A) APPLICATION NUMBER: US 08/089,372
 - (B) FILING DATE: 08-JUL-1993
- (vii) PRIOR APPLICATION DATA:
 - (A) APPLICATION NUMBER: AU PL4081/92
 - (B) FILING DATE: 14-AUG-1992
- (viii) ATTORNEY/AGENT INFORMATION:
 - (A) NAME: Trecartin, Richard F.
 - (B) REGISTRATION NUMBER: 31,801
 - (C) REFERENCE/DOCKET NUMBER: A-58631-3/RFT/DSS
- (ix) TELECOMMUNICATION INFORMATION:
 - (A) TELEPHONE: (415) 781-1989
 - (B) TELEFAX: (415) 398-3249
 - (C) TELEX: 910 277299

(2) INFORMATION FOR SEQ ID NO:1:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 13 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: unknown
 - (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

GGATCCACAG NNN

13

(2) INFORMATION FOR SEQ ID NO:2:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 28 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: unknown
 - (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

ATGGGCGATG CCGGCGTCGC GTTCACAG

28

(2) INFORMATION FOR SEQ ID NO:3:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 27 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: unknown
 - (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

ATGGAGGATG CTGGAGTGGC GTCACAG

27

(2) INFORMATION FOR SEQ ID NO:4:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 27 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: unknown
 - (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

ATGAGCGAGG CCGGCGTCGC GTCACAG

27

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 30 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: unknown
- (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

CCATCGATGC CGGACTGGTA TCCCAGGGGG

30

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 31 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: unknown
- (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

CCATCGATGC CGGACTGGTA TCCCGAGGGA C

31

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 39 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: unknown
- (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

CCATCGATGA TCCAGCCTCC TCGCGGCGCC GGATGGGCA

39

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 39 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: unknown
- (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

GCTCTAGATC CATTGCCCAT CCGAAGATGC CCATCCGGC

39

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 39 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: unknown
- (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

CCATCGATTT ATGCCGAGAA GGTAACCAGA GAAACACAC

39

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 39 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: unknown
- (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

GCTCTAGACC AGGTAATATA CCACAACGTG TGTTTCTCT

39

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 45 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: unknown
- (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

GGGGGAATT CATTTAGGTG AACTATAGT TCTGCCTCCC CGGAC

45

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 27 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: unknown
- (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

GGGGGGATCC TGGTATCCCA GGGGGGC

27

(2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 28 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: unknown
- (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

CCGGAAGCTT GTTTTTCTTT CTTTACCA

28

(2) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 46 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: unknown
- (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

GGGGGATCCG ATGGTATCCC GAGGGACGCT CAGCAGGTGG CATAGG

46

(2) INFORMATION FOR SEQ ID NO:15:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 52 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: unknown
- (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

AAATAATTTT GTTACTTTAG AAGGAGATAT ACATATGAGC GAGCGAGCAC AC

52

(2) INFORMATION FOR SEQ ID NO:16:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 55 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: unknown
- (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

AAATAATTTT GTTTAACCTT AAGAAGGAGA TCTACATATG CTGGAGTGGC GTCAC

55

(2) INFORMATION FOR SEQ ID NO:17:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 30 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: unknown
- (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

GGAGATCTAC ATATGGGAGA TGCTGGAGTG

30

(2) INFORMATION FOR SEQ ID NO:18:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 17 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: unknown
- (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

GTAGCGAACG TCGAGAA

17

(2) INFORMATION FOR SEQ ID NO:19:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 31 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: unknown
- (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

GGGGGATCCT CAGTTGTCAG TGGCGGGGTA G

31

(2) INFORMATION FOR SEQ ID NO:20:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 28 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: unknown

(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

GGGGATCCCT AATTGGCACG AGCGGCGC

28

(2) INFORMATION FOR SEQ ID NO:21:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 29 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: unknown
- (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

AATTACATAT GGCGGCCGCC GTTTCTGCC

29

(2) INFORMATION FOR SEQ ID NO:22:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 29 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: unknown
- (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

AATTACATAT GTTCGCGGCC GCCGTTTCT

29

(2) INFORMATION FOR SEQ ID NO:23:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 19 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: unknown
- (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

Phe Ala Ala Ala Val Ser Ala Phe Ala Ala Asn Met Leu Ser Ser Val
1 5 10 15

Leu Lys Ser

(2) INFORMATION FOR SEQ ID NO:24:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 20 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS: unknown
 - (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

Pro Thr Leu Val Asp Gln Gly Phe Trp Ile Gly Gly Gln Tyr Ala Leu
 1 5 10 15

Thr Pro Thr Ser
 20

(2) INFORMATION FOR SEQ ID NO:25:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 6 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS: unknown
 - (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

Phe Ala Ala Ala Val Ser
 1 5

(2) INFORMATION FOR SEQ ID NO:26:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 23 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: unknown
 - (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:

GCGCCCCCUG GGAUACCAGG AUC

23

(2) INFORMATION FOR SEQ ID NO:27:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 17 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: unknown
 - (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:

TCAGCAGGTG GCATAGG

17

(2) INFORMATION FOR SEQ ID NO:28:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 32 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: unknown
- (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: DNA

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 6..32

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:28:

CCCAT ATG GGC GAT GCC GGC GTC GCG TCA CAG
Met Gly Asp Ala Gly Val Ala Ser Gln
1 5

32

(2) INFORMATION FOR SEQ ID NO:29:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 9 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:29:

Met Gly Asp Ala Gly Val Ala Ser Gln
1 5

(2) INFORMATION FOR SEQ ID NO:30:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 32 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: unknown
- (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: DNA

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 6..32

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:30:

CCCAT ATG AGC GAG GCC GGC GTC GCG TCA CAG
Met Ser Glu Ala Gly Val Ala Ser Gln
1 5

32

(2) INFORMATION FOR SEQ ID NO:31:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 9 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:31:

Met Ser Glu Ala Gly Val Ala Ser Gln
1 5

(2) INFORMATION FOR SEQ ID NO:32:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 27 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: unknown
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: DNA

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 1..27

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:32:

ATG GGA GAT GCT GGA GTG GCG TCA CAG
Met Gly Asp Ala Gly Val Ala Ser Gln
1 5

27

(2) INFORMATION FOR SEQ ID NO:33:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 9 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:33:

Met Gly Asp Ala Gly Val Ala Ser Gln
1 5

(2) INFORMATION FOR SEQ ID NO:34:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 27 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: unknown
 - (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:34:

GGGGGATCCC GCGGATTTAT GAGCGAG

27

(2) INFORMATION FOR SEQ ID NO:35:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 32 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: unknown
 - (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:35:

GGGGGATCCC GCGGAGACAT GAGCGAGCAC AC

32

(2) INFORMATION FOR SEQ ID NO:36:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 34 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: unknown
 - (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:36:

GGGGGATCCA GCGACATGAG AGATGCTGGA GTGG

34

(2) INFORMATION FOR SEQ ID NO:37:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 34 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: unknown
 - (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:37:

GGGGGATCCA GCGACATGAG AGATGCTGGA GTGG

34

(2) INFORMATION FOR SEQ ID NO:38:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 26 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: unknown
 (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:38:

GGGGGATCCG TTCTGCCTCC CCGGAC

26

(2) INFORMATION FOR SEQ ID NO:39:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 5312 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: unknown
 (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: DNA

(ix) FEATURE:

- (A) NAME/KEY: CDS
 (B) LOCATION: 37..5148

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:39:

GTTCTGCCTC CCCC	GGACGG	TAAATATAGG	GGAACA	ATG	TAC	GCG	AAA	GCG	ACA		54
				Met	Tyr	Ala	Lys	Ala	Thr		
				1				5			
GAC GTG GCG CGT GTC TAC GCC GCG GCA GAT GTC GCC TAC GCG AAC GTA											102
Asp Val Ala Arg Val Tyr Ala Ala Ala Asp Val Ala Tyr Ala Asn Val											
	10		15					20			
CTG CAG CAG AGA GCA GTC AAG TTG GAC TTC GCC CCG CCA CTG AAG GCA											150
Leu Gln Gln Arg Ala Val Lys Leu Asp Phe Ala Pro Pro Leu Lys Ala											
	25		30					35			
CTA GAA ACC CTC CAC AGA CTG TAC TAT CCG CTG CGC TTC AAA GGG GGC											198
Leu Glu Thr Leu His Arg Leu Tyr Tyr Pro Leu Arg Phe Lys Gly Gly											
	40		45					50			
ACT TTA CCC CCG ACA CAA CAC CCG ATC CTG GCC GGG CAC CAA CGT GTC											246
Thr Leu Pro Pro Thr Gln His Pro Ile Leu Ala Gly His Gln Arg Val											
	55		60				65			70	
GCA GAA GAG GTT CTG CAC AAT TTC GCC AGG GGA CGT AGC ACA GTG CTC											294
Ala Glu Glu Val Leu His Asn Phe Ala Arg Gly Arg Ser Thr Val Leu											
	75		80					85			

GAG ATA GGG CCG TCT CTG CAC AGC GCA CTT AAG CTA CAT GGG GCA CCG	342
Glu Ile Gly Pro Ser Leu His Ser Ala Leu Lys Leu His Gly Ala Pro	
90 95 100	
AAC GCC CCC GTC GCA GAC TAT CAC GGG TGC ACC AAG TAC GGC ACC CGC	390
Asn Ala Pro Val Ala Asp Tyr His Gly Cys Thr Lys Tyr Gly Thr Arg	
105 110 115	
GAC GGC TCG CGA CAC ATT ACG GCC TTA GAG TCT AGA TCC GTC GCC ACA	438
Asp Gly Ser Arg His Ile Thr Ala Leu Glu Ser Arg Ser Val Ala Thr	
120 125 130	
GGC CGG CCC GAG TTC AAG GCC GAC GCC TCA CTG CTC GCC AAC GGC ATT	486
Gly Arg Pro Glu Phe Lys Ala Asp Ala Ser Leu Leu Ala Asn Gly Ile	
135 140 145 150	
GCC TCC CGC ACC TTC TGC GTC GAC GGA GTC GGC TCT TGC GCG TTC AAA	534
Ala Ser Arg Thr Phe Cys Val Asp Gly Val Gly Ser Cys Ala Phe Lys	
155 160 165	
TCG CGC GTT GGA ATT GCC AAT CAC TCC CTC TAT GAC GTG ACC CTA GAG	582
Ser Arg Val Gly Ile Ala Asn His Ser Leu Tyr Asp Val Thr Leu Glu	
170 175 180	
GAG CTG GCC AAT GCG TTT GAG AAC CAC GGA CTT CAC ATG GTC CGC GCG	630
Glu Leu Ala Asn Ala Phe Glu Asn His Gly Leu His Met Val Arg Ala	
185 190 195	
TTC ATG CAC ATG CCA GAA GAG CTG CTC TAC ATG GAC AAC GTG GTT AAT	678
Phe Met His Met Pro Glu Glu Leu Leu Tyr Met Asp Asn Val Val Asn	
200 205 210	
GCC GAG CTC GGC TAC CGC TTC CAC GTT ATT GAA GAG CCT ATG GCT GTG	726
Ala Glu Leu Gly Tyr Arg Phe His Val Ile Glu Glu Pro Met Ala Val	
215 220 225 230	
AAG GAC TGC GCA TTC CAG GGG GGG GAC CTC CGT CTC CAC TTC CCT GAG	774
Lys Asp Cys Ala Phe Gln Gly Gly Asp Leu Arg Leu His Phe Pro Glu	
235 240 245	
TTG GAC TTC ATC AAC GAG AGC CAA GAG CGG CGC ATC GAG AGG CTG GCC	822
Leu Asp Phe Ile Asn Glu Ser Gln Glu Arg Arg Ile Glu Arg Leu Ala	
250 255 260	
GCC CGC GGC TCC TAC TCC AGA CGC GCC GTC ATT TTC TCC GGC GAC GAC	870
Ala Arg Gly Ser Tyr Ser Arg Arg Ala Val Ile Phe Ser Gly Asp Asp	
265 270 275	
GAC TGG GGT GAT GCG TAC TTA CAC GAC TTC CAC ACA TGG CTC GCC TAC	918
Asp Trp Gly Asp Ala Tyr Leu His Asp Phe His Thr Trp Leu Ala Tyr	
280 285 290	
CTA CTG GTG AGG AAC TAC CCC ACT CCG TTT GGT TTC TCA CTC CAT ATA	966
Leu Leu Val Arg Asn Tyr Pro Thr Pro Phe Gly Phe Ser Leu His Ile	
295 300 305 310	

GAA	GTC	CAG	AGG	CGC	CAC	GGC	TCC	AGC	ATT	GAG	CTG	CGC	ATC	ACT	CGC	1014
Glu	Val	Gln	Arg	Arg	His	Gly	Ser	Ser	Ile	Glu	Leu	Arg	Ile	Thr	Arg	
				315					320						325	
GCG	CCA	CCT	GGA	GAC	CGC	ATG	CTG	GCC	GTC	GTC	CCA	AGG	ACG	TCC	CAA	1062
Ala	Pro	Pro	Gly	Asp	Arg	Met	Leu	Ala	Val	Val	Pro	Arg	Thr	Ser	Gln	
			330					335					340			
GGC	CTC	TGC	AGA	ATC	CCA	AAC	ATC	TTT	TAT	TAC	GCC	GAC	GCG	TCG	GGC	1110
Gly	Leu	Cys	Arg	Ile	Pro	Asn	Ile	Phe	Tyr	Tyr	Ala	Asp	Ala	Ser	Gly	
		345					350					355				
ACT	GAG	CAT	AAG	ACC	ATC	CTT	ACG	TCA	CAG	CAC	AAA	GTC	AAC	ATG	CTG	1158
Thr	Glu	His	Lys	Thr	Ile	Leu	Thr	Ser	Gln	His	Lys	Val	Asn	Met	Leu	
	360					365					370					
CTC	AAT	TTT	ATG	CAA	ACG	CGT	CCT	GAG	AAG	GAA	CTA	GTC	GAC	ATG	ACC	1206
Leu	Asn	Phe	Met	Gln	Thr	Arg	Pro	Glu	Lys	Glu	Leu	Val	Asp	Met	Thr	
375				380						385					390	
GTC	TTG	ATG	TCG	TTC	GCG	CGC	GCT	AGG	CTG	CGC	GCG	ATC	GTG	GTC	GCC	1254
Val	Leu	Met	Ser	Phe	Ala	Arg	Ala	Arg	Leu	Arg	Ala	Ile	Val	Val	Ala	
				395					400					405		
TCA	GAA	GTC	ACC	GAG	AGC	TCC	TGG	AAC	ATC	TCA	CCG	GCT	GAC	CTG	GTC	1302
Ser	Glu	Val	Thr	Glu	Ser	Ser	Trp	Asn	Ile	Ser	Pro	Ala	Asp	Leu	Val	
			410					415					420			
CGC	ACT	GTC	GTG	TCT	CTT	TAC	GTC	CTC	CAC	ATC	ATC	GAG	CGC	CGA	AGG	1350
Arg	Thr	Val	Val	Ser	Leu	Tyr	Val	Leu	His	Ile	Ile	Glu	Arg	Arg	Arg	
		425					430					435				
GCT	GCG	GTC	GCT	GTC	AAG	ACC	GCC	AAG	GAC	GAC	GTC	TTT	GGA	GAG	ACT	1398
Ala	Ala	Val	Ala	Val	Lys	Thr	Ala	Lys	Asp	Asp	GTC	Val	Phe	Gly	Glu	
	440					445					450					
TCG	TTC	TGG	GAG	AGT	CTC	AAG	CAC	GTC	TTG	GGC	TCC	TGT	TGC	GGT	CTG	1446
Ser	Phe	Trp	Glu	Ser	Leu	Lys	His	Val	Leu	Gly	Ser	Cys	Cys	Gly	Leu	
455					460					465					470	
CGC	AAC	CTC	AAA	GGC	ACC	GAC	GTC	GTC	TTT	ACT	AAG	CGC	GTC	GTC	GAT	1494
Arg	Asn	Leu	Lys	Gly	Thr	Asp	Val	Val	Phe	Thr	Lys	Arg	Val	Val	Asp	
				475					480					485		
AAG	TAC	CGA	GTC	CAC	TCG	CTC	GGA	GAC	ATA	ATC	TGC	GAC	GTC	CGC	CTG	1542
Lys	Tyr	Arg	Val	His	Ser	Leu	Gly	Asp	Ile	Ile	Cys	Asp	Val	Arg	Leu	
			490					495					500			
TCC	CCT	GAA	CAG	GTC	GGC	TTC	CTG	CCG	TCC	CGC	GTA	CCA	CCT	GCC	CGC	1590
Ser	Pro	Glu	Gln	Val	Gly	Phe	Leu	Pro	Ser	Arg	Val	Pro	Pro	Ala	Arg	
		505					510					515				
GTC	TTT	CAC	GAC	AGG	GAA	GAG	CTT	GAG	GTC	CTT	CGC	GAA	GCT	GGC	TGC	1638
Val	Phe	His	Asp	Arg	Glu	Glu	Leu	Glu	Val	Leu	Arg	Glu	Ala	Gly	Cys	
	520					525					530					

TAC	AAC	GAA	CGT	CCG	GTA	CCT	TCC	ACT	CCT	CCT	GTG	GAG	GAG	CCC	CAA	1686
Tyr	Asn	Glu	Arg	Pro	Val	Pro	Ser	Thr	Pro	Pro	Val	Glu	Glu	Pro	Gln	
535					540					545					550	
GGT	TTC	GAC	GCC	GAC	TTG	TGG	CAC	GCG	ACC	GCG	GCC	TCA	CTC	CCC	GAG	1734
Gly	Phe	Asp	Ala	Asp	Leu	Trp	His	Ala	Thr	Ala	Ala	Ser	Leu	Pro	Glu	
				555					560						565	
TAC	CGC	GCC	ACC	TTG	CAG	GCA	GGT	CTC	AAC	ACC	GAC	GTC	AAG	CAG	CTC	1782
Tyr	Arg	Ala	Thr	Leu	Gln	Ala	Gly	Leu	Asn	Thr	Asp	Val	Lys	Gln	Leu	
			570					575					580			
AAG	ATC	ACC	CTC	GAG	AAC	GCC	CTC	AAG	ACC	ATC	GAC	GGG	CTC	ACC	CTC	1830
Lys	Ile	Thr	Leu	Glu	Asn	Ala	Leu	Lys	Thr	Ile	Asp	Gly	Leu	Thr	Leu	
		585					590					595				
TCC	CCA	GTC	AGA	GGC	CTC	GAG	ATG	TAC	GAG	GGC	CCG	CCA	GGC	AGC	GGC	1878
Ser	Pro	Val	Arg	Gly	Leu	Glu	Met	Tyr	Glu	Gly	Pro	Pro	Gly	Ser	Gly	
	600					605					610					
AAG	ACG	GGC	ACC	CTC	ATC	GCC	GCC	CTT	GAG	GCC	GCG	GGC	GGT	AAA	GCA	1926
Lys	Thr	Gly	Thr	Leu	Ile	Ala	Ala	Leu	Glu	Ala	Ala	Gly	Gly	Lys	Ala	
615					620					625					630	
CTT	TAC	GTG	GCA	CCC	ACC	AGA	GAA	CTG	AGA	GAG	GCT	ATG	GAC	CGG	CGG	1974
Leu	Tyr	Val	Ala	Pro	Thr	Arg	Glu	Leu	Arg	Glu	Ala	Met	Asp	Arg	Arg	
				635					640					645		
ATC	AAA	CCG	CCG	TCC	GCC	TCG	GCT	ACG	CAA	CAT	GTC	GCC	CTT	GCG	ATT	2022
Ile	Lys	Pro	Pro	Ser	Ala	Ser	Ala	Thr	Gln	His	Val	Ala	Leu	Ala	Ile	
			650					655					660			
CTC	CGT	CGT	GCC	ACC	GCC	GAG	GGC	GCC	CCT	TTC	GCT	ACC	GTG	GTT	ATC	2070
Leu	Arg	Arg	Ala	Thr	Ala	Glu	Gly	Ala	Pro	Phe	Ala	Thr	Val	Val	Ile	
		665					670					675				
GAC	GAG	TGC	TTC	ATG	TTC	CCG	CTC	GTG	TAC	GTC	GCG	ATC	GTG	CAC	GCC	2118
Asp	Glu	Cys	Phe	Met	Phe	Pro	Leu	Val	Tyr	Val	Ala	Ile	Val	His	Ala	
	680					685					690					
TTG	TCC	CCG	AGC	TCA	CGA	ATA	GTC	CTT	GTA	GGG	GAC	GTC	CAC	CAA	ATC	2166
Leu	Ser	Pro	Ser	Ser	Arg	Ile	Val	Leu	Val	Gly	Asp	Val	His	Gln	Ile	
695					700					705					710	
GGG	TTT	ATA	GAC	TTC	CAA	GGC	ACA	AGC	GCG	AAC	ATG	CCG	CTC	GTT	CGC	2214
Gly	Phe	Ile	Asp	Phe	Gln	Gly	Thr	Ser	Ala	Asn	Met	Pro	Leu	Val	Arg	
				715					720					725		
GAC	GTC	GTT	AAG	CAG	TGC	CGT	CGG	CGC	ACT	TTC	AAC	CAA	ACC	AAG	CGC	2262
Asp	Val	Val	Lys	Gln	Cys	Arg	Arg	Arg	Thr	Phe	Asn	Gln	Thr	Lys	Arg	
			730					735					740			
TGT	CCG	GCC	GAC	GTC	GTT	GCC	ACC	ACG	TTT	TTC	CAG	AGC	TTG	TAC	CCC	2310
Cys	Pro	Ala	Asp	Val	Val	Ala	Thr	Thr	Phe	Phe	Gln	Ser	Leu	Tyr	Pro	
		745					750						755			

GGG TGC ACA ACC ACC TCA GGG TGC GTC GCA TCC ATC AGC CAC GTC GCC	2358
Gly Cys Thr Thr Thr Ser Gly Cys Val Ala Ser Ile Ser His Val Ala	
760 765 770	
CCA GAC TAC CGC AAC AGC CAG GCG CAA ACG CTC TGC TTC ACG CAG GAG	2406
Pro Asp Tyr Arg Asn Ser Gln Ala Gln Thr Leu Cys Phe Thr Gln Glu	
775 780 785 790	
GAA AAG TCG CGC CAC GGG GCT GAG GGC GCG ATG ACT GTG CAC GAA GCG	2454
Glu Lys Ser Arg His Gly Ala Glu Gly Ala Met Thr Val His Glu Ala	
795 800 805	
CAG GGA CGC ACT TTT GCG TCT GTC ATT CTG CAT TAC AAC GGC TCC ACA	2502
Gln Gly Arg Thr Phe Ala Ser Val Ile Leu His Tyr Asn Gly Ser Thr	
810 815 820	
GCA GAG CAG AAG CTC CTC GCT GAG AAG TCG CAC CTT CTA GTC GGC ATC	2550
Ala Glu Gln Lys Leu Leu Ala Glu Lys Ser His Leu Leu Val Gly Ile	
825 830 835	
ACG CGC CAC ACC AAC CAC CTG TAC ATC CGC GAC CCG ACA GGT GAC ATT	2598
Thr Arg His Thr Asn His Leu Tyr Ile Arg Asp Pro Thr Gly Asp Ile	
840 845 850	
GAG AGA CAA CTC AAC CAT AGC GCG AAA GCC GAG GTG TTT ACA GAC ATC	2646
Glu Arg Gln Leu Asn His Ser Ala Lys Ala Glu Val Phe Thr Asp Ile	
855 860 865 870	
CCT GCA CCC CTG GAG ATC ACG ACT GTC AAA CCG AGT GAA GAG GTG CAG	2694
Pro Ala Pro Leu Glu Ile Thr Thr Val Lys Pro Ser Glu Glu Val Gln	
875 880 885	
CGC AAC GAA GTG ATG GCA ACG ATA CCC CCG CAG AGT GCC ACG CCG CAC	2742
Arg Asn Glu Val Met Ala Thr Ile Pro Pro Gln Ser Ala Thr Pro His	
890 895 900	
GGA GCA ATC CAT CTG CTC CGC AAG AAC TTC GGG GAC CAA CCC GAC TGT	2790
Gly Ala Ile His Leu Leu Arg Lys Asn Phe Gly Asp Gln Pro Asp Cys	
905 910 915	
GGC TGT GTC GCT TTG GCG AAG ACC GGC TAC GAG GTG TTT GGC GGT CGT	2838
Gly Cys Val Ala Leu Ala Lys Thr Gly Tyr Glu Val Phe Gly Gly Arg	
920 925 930	
GCC AAA ATC AAC GTA GAG CTT GCC GAA CCC GAC GCG ACC CCG AAG CCG	2886
Ala Lys Ile Asn Val Glu Leu Ala Glu Pro Asp Ala Thr Pro Lys Pro	
935 940 945 950	
CAT AGG GCG TTC CAG GAA GGG GTA CAG TGG GTC AAG GTC ACC AAC GCG	2934
His Arg Ala Phe Gln Glu Gly Val Gln Trp Val Lys Val Thr Asn Ala	
955 960 965	
TCT AAC AAA CAC CAG GCG CTC CAG ACG CTG TTG TCC CGC TAC ACC AAG	2982
Ser Asn Lys His Gln Ala Leu Gln Thr Leu Leu Ser Arg Tyr Thr Lys	
970 975 980	

CGA AGC GCT GAC CTG CCG CTA CAC GAA GCT AAG GAG GAC GTC AAA CGC	3030
Arg Ser Ala Asp Leu Pro Leu His Glu Ala Lys Glu Asp Val Lys Arg	
985 990 995	
ATG CTA AAC TCG CTT GAC CGA CAT TGG GAC TGG ACT GTC ACT GAA GAC	3078
Met Leu Asn Ser Leu Asp Arg His Trp Asp Trp Thr Val Thr Glu Asp	
1000 1005 1010	
GCC CGT GAC CGA GCT GTC TTC GAG ACC CAG CTC AAG TTC ACC CAA CGC	3126
Ala Arg Asp Arg Ala Val Phe Glu Thr Gln Leu Lys Phe Thr Gln Arg	
1015 1020 1025 1030	
GGC GGC ACC GTC GAA GAC CTG CTG GAG CCA GAC GAC CCC TAC ATC CGT	3174
Gly Gly Thr Val Glu Asp Leu Leu Glu Pro Asp Asp Pro Tyr Ile Arg	
1035 1040 1045	
GAC ATA GAC TTC CTT ATG AAG ACT CAG CAG AAA GTG TCG CCC AAG CCG	3222
Asp Ile Asp Phe Leu Met Lys Thr Gln Gln Lys Val Ser Pro Lys Pro	
1050 1055 1060	
ATC AAT ACG GGC AAG GTC GGG CAG GGG ATC GCC GCT CAC TCA AAG TCT	3270
Ile Asn Thr Gly Lys Val Gly Gln Gly Ile Ala Ala His Ser Lys Ser	
1065 1070 1075	
CTC AAC TTC GTC CTC GCC GCT TGG ATA CGC ATA CTC GAG GAG ATA CTC	3318
Leu Asn Phe Val Leu Ala Ala Trp Ile Arg Ile Leu Glu Glu Ile Leu	
1080 1085 1090	
CGT ACC GGG AGC CGC ACG GTC CGG TAC AGC AAC GGT CTC CCC GAC GAA	3366
Arg Thr Gly Ser Arg Thr Val Arg Tyr Ser Asn Gly Leu Pro Asp Glu	
1095 1100 1105 1110	
GAA GAG GCC ATG CTG CTC GAA GCG AAG ATC AAT CAA GTC CCA CAC GCC	3414
Glu Glu Ala Met Leu Leu Glu Ala Lys Ile Asn Gln Val Pro His Ala	
1115 1120 1125	
ACG TTC GTC TCG GCG GAC TGG ACC GAG TTT GAC ACC GCC CAC AAT AAC	3462
Thr Phe Val Ser Ala Asp Trp Thr Glu Phe Asp Thr Ala His Asn Asn	
1130 1135 1140	
ACG AGT GAG CTG CTC TTC GCC GCC CTT TTA GAG CGC ATC GGC ACG CCT	3510
Thr Ser Glu Leu Leu Phe Ala Ala Leu Leu Glu Arg Ile Gly Thr Pro	
1145 1150 1155	
GCA GCT GCC GTT AAT CTA TTC AGA GAA CGG TGT GGG AAA CGC ACC TTG	3558
Ala Ala Ala Val Asn Leu Phe Arg Glu Arg Cys Gly Lys Arg Thr Leu	
1160 1165 1170	
CGA GCG AAG GGT CTA GGC TCC GTT GAA GTC GAC GGT CTG CTC GAC TCC	3606
Arg Ala Lys Gly Leu Gly Ser Val Glu Val Asp Gly Leu Leu Asp Ser	
1175 1180 1185 1190	
GGC GCA GCT TGG ACG CCT TGC CGC AAC ACC ATC TTC TCT GCC GCC GTC	3654
Gly Ala Ala Trp Thr Pro Cys Arg Asn Thr Ile Phe Ser Ala Ala Val	
1195 1200 1205	

ATG CTC ACG CTC TTC CGC GGC GTC AAG TTC GCA GCT TTC AAA GGC GAC	3702
Met Leu Thr Leu Phe Arg Gly Val Lys Phe Ala Ala Phe Lys Gly Asp	
1210 1215 1220	
GAC TCG CTC CTC TGT GGT AGC CAT TAC CTC CGT TTC GAC GCT AGC CGC	3750
Asp Ser Leu Leu Cys Gly Ser His Tyr Leu Arg Phe Asp Ala Ser Arg	
1225 1230 1235	
CTT CAC ATG GGC GAA CGT TAC AAG ACC AAA CAT TTG AAG GTC GAG GTG	3798
Leu His Met Gly Glu Arg Tyr Lys Thr Lys His Leu Lys Val Glu Val	
1240 1245 1250	
CAG AAA ATC GTG CCG TAC ATC GGA CTC CTC GTC TCC GCT GAG CAG GTC	3846
Gln Lys Ile Val Pro Tyr Ile Gly Leu Leu Val Ser Ala Glu Gln Val	
1255 1260 1265 1270	
GTC CTC GAC CCT GTC AGG AGC GCT CTC AAG ATA TTT GGG CGC TGC TAC	3894
Val Leu Asp Pro Val Arg Ser Ala Leu Lys Ile Phe Gly Arg Cys Tyr	
1275 1280 1285	
ACA AGC GAA CTC CTT TAC TCC AAG TAC GTG GAG GCT GTG AGA GAC ATC	3942
Thr Ser Glu Leu Leu Tyr Ser Lys Tyr Val Glu Ala Val Arg Asp Ile	
1290 1295 1300	
ACC AAG GGC TGG AGT GAC GCC CGC TAC CAC AGC CTC CTG TGC CAC ATG	3990
Thr Lys Gly Trp Ser Asp Ala Arg Tyr His Ser Leu Leu Cys His Met	
1305 1310 1315	
TCA GCA TGC TAC TAC AAT TAC GCG CCG GAG TCT GCG GCG TAC ATC ATC	4038
Ser Ala Cys Tyr Tyr Asn Tyr Ala Pro Glu Ser Ala Ala Tyr Ile Ile	
1320 1325 1330	
GAC GCT GTT GTT CGC TTT GGG CGC GGC GAC TTC CCG TTT GAA CAA CTG	4086
Asp Ala Val Val Arg Phe Gly Arg Gly Asp Phe Pro Phe Glu Gln Leu	
1335 1340 1345 1350	
CGC GTG GTG CGT GCC CAT GTG CAG GCA CCC GAC GCT TAC AGC AGC ACG	4134
Arg Val Val Arg Ala His Val Gln Ala Pro Asp Ala Tyr Ser Ser Thr	
1355 1360 1365	
TAT CCG GCT AAC GTG CGC GCA TCG TGC CTT GAC CAC GTC TTC GAG CCC	4182
Tyr Pro Ala Asn Val Arg Ala Ser Cys Leu Asp His Val Phe Glu Pro	
1370 1375 1380	
CGC CAG GCC GCC GCC CCG GCA GGT TTC GTT GCG ACA TGT GCG AAG CCG	4230
Arg Gln Ala Ala Ala Pro Ala Gly Phe Val Ala Thr Cys Ala Lys Pro	
1385 1390 1395	
GAA ACG CCT TCT TCA CTT ACC GCG AAA GCT GGT GTT TCT GCG ACT ACA	4278
Glu Thr Pro Ser Ser Leu Thr Ala Lys Ala Gly Val Ser Ala Thr Thr	
1400 1405 1410	
AGC CAC GTT GCG ACT GGG ACT GCG CCC CCG GAG TCT CCA TGG GAT GCA	4326
Ser His Val Ala Thr Gly Thr Ala Pro Pro Glu Ser Pro Trp Asp Ala	
1415 1420 1425 1430	

CCT GCA GCC AAC AGC TTT TCG GAG TTA TTG ACA CCG GAG ACC CCG TCC	4374
Pro Ala Ala Asn Ser Phe Ser Glu Leu Leu Thr Pro Glu Thr Pro Ser	
1435 1440 1445	
ACA TCA TCC TCG CCG TCA TCG TCT TCA TCG GAC TCC TCT ACA TCG TGT	4422
Thr Ser Ser Ser Pro Ser Ser Ser Ser Ser Asp Ser Ser Thr Ser Cys	
1450 1455 1460	
GGA AGG TCG CTC AGT GGT GGA GAC ACC GCA AGG ACC ACA GAA GAC TTG	4470
Gly Arg Ser Leu Ser Gly Gly Asp Thr Ala Arg Thr Thr Glu Asp Leu	
1465 1470 1475	
AAC AGC AGA AAG CCG CCT TCG CAA GAC AGG CAA TCA CGC TCG TCT GAA	4518
Asn Ser Arg Lys Pro Pro Ser Gln Asp Arg Gln Ser Arg Ser Ser Glu	
1480 1485 1490	
TGT CTG GAC AGA AGC GGA GAA AGG ACA GGC AGT TCG TTA ACT GCC CCC	4566
Cys Leu Asp Arg Ser Gly Glu Arg Thr Gly Ser Ser Leu Thr Ala Pro	
1495 1500 1505 1510	
ACT GCT CCG AGC CCC TCA TTC TCA TTT TCG GAA AGA GCT CGA CTG GCG	4614
Thr Ala Pro Ser Pro Ser Phe Ser Phe Ser Glu Arg Ala Arg Leu Ala	
1515 1520 1525	
ACC GGG CCG ACT GTC GCC GCT GCG ACA TCA CCT TCG GCA ACC CCA TCC	4662
Thr Gly Pro Thr Val Ala Ala Ala Thr Ser Pro Ser Ala Thr Pro Ser	
1530 1535 1540	
TGC GCC ACG GAC CAG GTT GCC GCG AGG ACC ACG CCG GAC TTT GCG CCT	4710
Cys Ala Thr Asp Gln Val Ala Ala Arg Thr Thr Pro Asp Phe Ala Pro	
1545 1550 1555	
TTC CTG GGT TCC CAG TCT GCC CGT GCT GTC TCG AAG CCG TAC CGG CCC	4758
Phe Leu Gly Ser Gln Ser Ala Arg Ala Val Ser Lys Pro Tyr Arg Pro	
1560 1565 1570	
CCC ACG ACT GCC CGT TGG AAA GAA GTC ACC CCG CTC CAC GCG TGG AAG	4806
Pro Thr Thr Ala Arg Trp Lys Glu Val Thr Pro Leu His Ala Trp Lys	
1575 1580 1585 1590	
GGC GTG ACC GGA GAC CGA CCG GAA GTC AGG GAG GAC CCG GAG ACA GCG	4854
Gly Val Thr Gly Asp Arg Pro Glu Val Arg Glu Asp Pro Glu Thr Ala	
1595 1600 1605	
GCG GTC GTC CAG GCT CTG ATC AGC GGC CGT TAT CCT CAG AAG ACG AAG	4902
Ala Val Val Gln Ala Leu Ile Ser Gly Arg Tyr Pro Gln Lys Thr Lys	
1610 1615 1620	
CTT TCC TCC GAC GCA TCC AAA GGC TAC TCA AGA ACT AAG GGA TGC TCA	4950
Leu Ser Ser Asp Ala Ser Lys Gly Tyr Ser Arg Thr Lys Gly Cys Ser	
1625 1630 1635	
CAA TCC ACC TCT TTT CCT GCC CCG AGT GCG GAT TAC CAG GCC CGC GAC	4998
Gln Ser Thr Ser Phe Pro Ala Pro Ser Ala Asp Tyr Gln Ala Arg Asp	
1640 1645 1650	

TGC CAG ACA GTC CGA GTC TGC CGC GCC GCT GCA GAG ATG GCG CGC TCA	5046
Cys Gln Thr Val Arg Val Cys Arg Ala Ala Glu Met Ala Arg Ser	
1655 1660 1665 1670	
TGT ATT CAC GAG CCG TTG GCT TCA TCT GCC GCC AGT GCC GAC TTG AAG	5094
Cys Ile His Glu Pro Leu Ala Ser Ser Ala Ala Ser Ala Asp Leu Lys	
1675 1680 1685	
CGC ATA CGC TCT ACC TCG GAC TCT GTT CCC GAT GTA AAG ATC AGC AAG	5142
Arg Ile Arg Ser Thr Ser Asp Ser Val Pro Asp Val Lys Ile Ser Lys	
1690 1695 1700	
AGC GCA TGAAGGAACA AAATTAGTTT CCTTGTTTCGT AAACAAGGTG GTCCCTCCCA	5198
Ser Ala	
TTGAGGTAAA GACTCTGGTG AGTCCTCAAC GTTACTCGTT GAGTCTGCTG CGGTTCGATT	5258
CCATTCCCAA GCAGCAAAGG GTGCGCAACT AGTACGGCGC CCCCTGGGAT ACCA	5312

(2) INFORMATION FOR SEQ ID NO:40:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1704 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:40:

Met Tyr Ala Lys Ala Thr Asp Val Ala Arg Val Tyr Ala Ala Ala Asp	
1 5 10 15	
Val Ala Tyr Ala Asn Val Leu Gln Gln Arg Ala Val Lys Leu Asp Phe	
20 25 30	
Ala Pro Pro Leu Lys Ala Leu Glu Thr Leu His Arg Leu Tyr Tyr Pro	
35 40 45	
Leu Arg Phe Lys Gly Gly Thr Leu Pro Pro Thr Gln His Pro Ile Leu	
50 55 60	
Ala Gly His Gln Arg Val Ala Glu Glu Val Leu His Asn Phe Ala Arg	
65 70 75 80	
Gly Arg Ser Thr Val Leu Glu Ile Gly Pro Ser Leu His Ser Ala Leu	
85 90 95	
Lys Leu His Gly Ala Pro Asn Ala Pro Val Ala Asp Tyr His Gly Cys	
100 105 110	
Thr Lys Tyr Gly Thr Arg Asp Gly Ser Arg His Ile Thr Ala Leu Glu	
115 120 125	

Ser Arg Ser Val Ala Thr Gly Arg Pro Glu Phe Lys Ala Asp Ala Ser
 130 135 140
 Leu Leu Ala Asn Gly Ile Ala Ser Arg Thr Phe Cys Val Asp Gly Val
 145 150 155 160
 Gly Ser Cys Ala Phe Lys Ser Arg Val Gly Ile Ala Asn His Ser Leu
 165 170 175
 Tyr Asp Val Thr Leu Glu Glu Leu Ala Asn Ala Phe Glu Asn His Gly
 180 185 190
 Leu His Met Val Arg Ala Phe Met His Met Pro Glu Glu Leu Leu Tyr
 195 200 205
 Met Asp Asn Val Val Asn Ala Glu Leu Gly Tyr Arg Phe His Val Ile
 210 215 220
 Glu Glu Pro Met Ala Val Lys Asp Cys Ala Phe Gln Gly Gly Asp Leu
 225 230 235 240
 Arg Leu His Phe Pro Glu Leu Asp Phe Ile Asn Glu Ser Gln Glu Arg
 245 250 255
 Arg Ile Glu Arg Leu Ala Ala Arg Gly Ser Tyr Ser Arg Arg Ala Val
 260 265 270
 Ile Phe Ser Gly Asp Asp Asp Trp Gly Asp Ala Tyr Leu His Asp Phe
 275 280 285
 His Thr Trp Leu Ala Tyr Leu Leu Val Arg Asn Tyr Pro Thr Pro Phe
 290 295 300
 Gly Phe Ser Leu His Ile Glu Val Gln Arg Arg His Gly Ser Ser Ile
 305 310 315 320
 Glu Leu Arg Ile Thr Arg Ala Pro Pro Gly Asp Arg Met Leu Ala Val
 325 330 335
 Val Pro Arg Thr Ser Gln Gly Leu Cys Arg Ile Pro Asn Ile Phe Tyr
 340 345 350
 Tyr Ala Asp Ala Ser Gly Thr Glu His Lys Thr Ile Leu Thr Ser Gln
 355 360 365
 His Lys Val Asn Met Leu Leu Asn Phe Met Gln Thr Arg Pro Glu Lys
 370 375 380
 Glu Leu Val Asp Met Thr Val Leu Met Ser Phe Ala Arg Ala Arg Leu
 385 390 395 400
 Arg Ala Ile Val Val Ala Ser Glu Val Thr Glu Ser Ser Trp Asn Ile
 405 410 415
 Ser Pro Ala Asp Leu Val Arg Thr Val Val Ser Leu Tyr Val Leu His
 420 425 430

Ile Ile Glu Arg Arg Arg Ala Ala Val Ala Val Lys Thr Ala Lys Asp
 435 440 445
 Asp Val Phe Gly Glu Thr Ser Phe Trp Glu Ser Leu Lys His Val Leu
 450 455 460
 Gly Ser Cys Cys Gly Leu Arg Asn Leu Lys Gly Thr Asp Val Val Phe
 465 470 475 480
 Thr Lys Arg Val Val Asp Lys Tyr Arg Val His Ser Leu Gly Asp Ile
 485 490 495
 Ile Cys Asp Val Arg Leu Ser Pro Glu Gln Val Gly Phe Leu Pro Ser
 500 505 510
 Arg Val Pro Pro Ala Arg Val Phe His Asp Arg Glu Glu Leu Glu Val
 515 520 525
 Leu Arg Glu Ala Gly Cys Tyr Asn Glu Arg Pro Val Pro Ser Thr Pro
 530 535 540
 Pro Val Glu Glu Pro Gln Gly Phe Asp Ala Asp Leu Trp His Ala Thr
 545 550 555 560
 Ala Ala Ser Leu Pro Glu Tyr Arg Ala Thr Leu Gln Ala Gly Leu Asn
 565 570 575
 Thr Asp Val Lys Gln Leu Lys Ile Thr Leu Glu Asn Ala Leu Lys Thr
 580 585 590
 Ile Asp Gly Leu Thr Leu Ser Pro Val Arg Gly Leu Glu Met Tyr Glu
 595 600 605
 Gly Pro Pro Gly Ser Gly Lys Thr Gly Thr Leu Ile Ala Ala Leu Glu
 610 615 620
 Ala Ala Gly Gly Lys Ala Leu Tyr Val Ala Pro Thr Arg Glu Leu Arg
 625 630 635 640
 Glu Ala Met Asp Arg Arg Ile Lys Pro Pro Ser Ala Ser Ala Thr Gln
 645 650 655
 His Val Ala Leu Ala Ile Leu Arg Arg Ala Thr Ala Glu Gly Ala Pro
 660 665 670
 Phe Ala Thr Val Val Ile Asp Glu Cys Phe Met Phe Pro Leu Val Tyr
 675 680 685
 Val Ala Ile Val His Ala Leu Ser Pro Ser Ser Arg Ile Val Leu Val
 690 695 700
 Gly Asp Val His Gln Ile Gly Phe Ile Asp Phe Gln Gly Thr Ser Ala
 705 710 715 720
 Asn Met Pro Leu Val Arg Asp Val Val Lys Gln Cys Arg Arg Arg Thr
 725 730 735

Phe	Asn	Gln	Thr	Lys	Arg	Cys	Pro	Ala	Asp	Val	Val	Ala	Thr	Thr	Phe	740	745	750
Phe	Gln	Ser	Leu	Tyr	Pro	Gly	Cys	Thr	Thr	Thr	Ser	Gly	Cys	Val	Ala	755	760	765
Ser	Ile	Ser	His	Val	Ala	Pro	Asp	Tyr	Arg	Asn	Ser	Gln	Ala	Gln	Thr	770	775	780
Leu	Cys	Phe	Thr	Gln	Glu	Glu	Lys	Ser	Arg	His	Gly	Ala	Glu	Gly	Ala	785	790	795
Met	Thr	Val	His	Glu	Ala	Gln	Gly	Arg	Thr	Phe	Ala	Ser	Val	Ile	Leu	805	810	815
His	Tyr	Asn	Gly	Ser	Thr	Ala	Glu	Gln	Lys	Leu	Leu	Ala	Glu	Lys	Ser	820	825	830
His	Leu	Leu	Val	Gly	Ile	Thr	Arg	His	Thr	Asn	His	Leu	Tyr	Ile	Arg	835	840	845
Asp	Pro	Thr	Gly	Asp	Ile	Glu	Arg	Gln	Leu	Asn	His	Ser	Ala	Lys	Ala	850	855	860
Glu	Val	Phe	Thr	Asp	Ile	Pro	Ala	Pro	Leu	Glu	Ile	Thr	Thr	Val	Lys	865	870	875
Pro	Ser	Glu	Glu	Val	Gln	Arg	Asn	Glu	Val	Met	Ala	Thr	Ile	Pro	Pro	885	890	895
Gln	Ser	Ala	Thr	Pro	His	Gly	Ala	Ile	His	Leu	Leu	Arg	Lys	Asn	Phe	900	905	910
Gly	Asp	Gln	Pro	Asp	Cys	Gly	Cys	Val	Ala	Leu	Ala	Lys	Thr	Gly	Tyr	915	920	925
Glu	Val	Phe	Gly	Gly	Arg	Ala	Lys	Ile	Asn	Val	Glu	Leu	Ala	Glu	Pro	930	935	940
Asp	Ala	Thr	Pro	Lys	Pro	His	Arg	Ala	Phe	Gln	Glu	Gly	Val	Gln	Trp	945	950	955
Val	Lys	Val	Thr	Asn	Ala	Ser	Asn	Lys	His	Gln	Ala	Leu	Gln	Thr	Leu	965	970	975
Leu	Ser	Arg	Tyr	Thr	Lys	Arg	Ser	Ala	Asp	Leu	Pro	Leu	His	Glu	Ala	980	985	990
Lys	Glu	Asp	Val	Lys	Arg	Met	Leu	Asn	Ser	Leu	Asp	Arg	His	Trp	Asp	995	1000	1005
Trp	Thr	Val	Thr	Glu	Asp	Ala	Arg	Asp	Arg	Ala	Val	Phe	Glu	Thr	Gln	1010	1015	1020
Leu	Lys	Phe	Thr	Gln	Arg	Gly	Gly	Thr	Val	Glu	Asp	Leu	Leu	Glu	Pro	1025	1030	1035

Asp	Asp	Pro	Tyr	Ile	Arg	Asp	Ile	Asp	Phe	Leu	Met	Lys	Thr	Gln	Gln	1045	1050	1055
Lys	Val	Ser	Pro	Lys	Pro	Ile	Asn	Thr	Gly	Lys	Val	Gly	Gln	Gly	Ile	1060	1065	1070
Ala	Ala	His	Ser	Lys	Ser	Leu	Asn	Phe	Val	Leu	Ala	Ala	Trp	Ile	Arg	1075	1080	1085
Ile	Leu	Glu	Glu	Ile	Leu	Arg	Thr	Gly	Ser	Arg	Thr	Val	Arg	Tyr	Ser	1090	1095	1100
Asn	Gly	Leu	Pro	Asp	Glu	Glu	Glu	Ala	Met	Leu	Leu	Glu	Ala	Lys	Ile	1105	1110	1115
Asn	Gln	Val	Pro	His	Ala	Thr	Phe	Val	Ser	Ala	Asp	Trp	Thr	Glu	Phe	1125	1130	1135
Asp	Thr	Ala	His	Asn	Asn	Thr	Ser	Glu	Leu	Leu	Phe	Ala	Ala	Leu	Leu	1140	1145	1150
Glu	Arg	Ile	Gly	Thr	Pro	Ala	Ala	Ala	Val	Asn	Leu	Phe	Arg	Glu	Arg	1155	1160	1165
Cys	Gly	Lys	Arg	Thr	Leu	Arg	Ala	Lys	Gly	Leu	Gly	Ser	Val	Glu	Val	1170	1175	1180
Asp	Gly	Leu	Leu	Asp	Ser	Gly	Ala	Ala	Trp	Thr	Pro	Cys	Arg	Asn	Thr	1185	1190	1195
Ile	Phe	Ser	Ala	Ala	Val	Met	Leu	Thr	Leu	Phe	Arg	Gly	Val	Lys	Phe	1205	1210	1215
Ala	Ala	Phe	Lys	Gly	Asp	Asp	Ser	Leu	Leu	Cys	Gly	Ser	His	Tyr	Leu	1220	1225	1230
Arg	Phe	Asp	Ala	Ser	Arg	Leu	His	Met	Gly	Glu	Arg	Tyr	Lys	Thr	Lys	1235	1240	1245
His	Leu	Lys	Val	Glu	Val	Gln	Lys	Ile	Val	Pro	Tyr	Ile	Gly	Leu	Leu	1250	1255	1260
Val	Ser	Ala	Glu	Gln	Val	Val	Leu	Asp	Pro	Val	Arg	Ser	Ala	Leu	Lys	1265	1270	1275
Ile	Phe	Gly	Arg	Cys	Tyr	Thr	Ser	Glu	Leu	Leu	Tyr	Ser	Lys	Tyr	Val	1285	1290	1295
Glu	Ala	Val	Arg	Asp	Ile	Thr	Lys	Gly	Trp	Ser	Asp	Ala	Arg	Tyr	His	1300	1305	1310
Ser	Leu	Leu	Cys	His	Met	Ser	Ala	Cys	Tyr	Tyr	Asn	Tyr	Ala	Pro	Glu	1315	1320	1325
Ser	Ala	Ala	Tyr	Ile	Ile	Asp	Ala	Val	Val	Arg	Phe	Gly	Arg	Gly	Asp	1330	1335	1340

Phe Pro Phe Glu Gln Leu Arg Val Val Arg Ala His Val Gln Ala Pro
 1345 1350 1355 1360
 Asp Ala Tyr Ser Ser Thr Tyr Pro Ala Asn Val Arg Ala Ser Cys Leu
 1365 1370 1375
 Asp His Val Phe Glu Pro Arg Gln Ala Ala Ala Pro Ala Gly Phe Val
 1380 1385 1390
 Ala Thr Cys Ala Lys Pro Glu Thr Pro Ser Ser Leu Thr Ala Lys Ala
 1395 1400 1405
 Gly Val Ser Ala Thr Thr Ser His Val Ala Thr Gly Thr Ala Pro Pro
 1410 1415 1420
 Glu Ser Pro Trp Asp Ala Pro Ala Ala Asn Ser Phe Ser Glu Leu Leu
 1425 1430 1435 1440
 Thr Pro Glu Thr Pro Ser Thr Ser Ser Ser Pro Ser Ser Ser Ser
 1445 1450 1455
 Asp Ser Ser Thr Ser Cys Gly Arg Ser Leu Ser Gly Gly Asp Thr Ala
 1460 1465 1470
 Arg Thr Thr Glu Asp Leu Asn Ser Arg Lys Pro Pro Ser Gln Asp Arg
 1475 1480 1485
 Gln Ser Arg Ser Ser Glu Cys Leu Asp Arg Ser Gly Glu Arg Thr Gly
 1490 1495 1500
 Ser Ser Leu Thr Ala Pro Thr Ala Pro Ser Pro Ser Phe Ser Phe Ser
 1505 1510 1515 1520
 Glu Arg Ala Arg Leu Ala Thr Gly Pro Thr Val Ala Ala Ala Thr Ser
 1525 1530 1535
 Pro Ser Ala Thr Pro Ser Cys Ala Thr Asp Gln Val Ala Ala Arg Thr
 1540 1545 1550
 Thr Pro Asp Phe Ala Pro Phe Leu Gly Ser Gln Ser Ala Arg Ala Val
 1555 1560 1565
 Ser Lys Pro Tyr Arg Pro Pro Thr Thr Ala Arg Trp Lys Glu Val Thr
 1570 1575 1580
 Pro Leu His Ala Trp Lys Gly Val Thr Gly Asp Arg Pro Glu Val Arg
 1585 1590 1595 1600
 Glu Asp Pro Glu Thr Ala Ala Val Val Gln Ala Leu Ile Ser Gly Arg
 1605 1610 1615
 Tyr Pro Gln Lys Thr Lys Leu Ser Ser Asp Ala Ser Lys Gly Tyr Ser
 1620 1625 1630
 Arg Thr Lys Gly Cys Ser Gln Ser Thr Ser Phe Pro Ala Pro Ser Ala
 1635 1640 1645

Asp Tyr Gln Ala Arg Asp Cys Gln Thr Val Arg Val Cys Arg Ala Ala
 1650 1655 1660

Ala Glu Met Ala Arg Ser Cys Ile His Glu Pro Leu Ala Ser Ser Ala
 1665 1670 1675 1680

Ala Ser Ala Asp Leu Lys Arg Ile Arg Ser Thr Ser Asp Ser Val Pro
 1685 1690 1695

Asp Val Lys Ile Ser Lys Ser Ala
 1700

(2) INFORMATION FOR SEQ ID NO:41:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 5312 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: unknown
- (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: DNA

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 4218..4514

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:41:

GTTCTGCCTC CCCCGGACGG TAAATATAGG GGAACAATGT ACGCGAAAGC GACAGACGTG	60
GCGCGTGTCT ACGCCGCGGC AGATGTCGCC TACGCGAACG TACTGCAGCA GAGAGCAGTC	120
AAGTTGGACT TCGCCCCGCC ACTGAAGGCA CTAGAAACCC TCCACAGACT GTACTATCCG	180
CTGCGCTTCA AAGGGGGCAC TTTACCCCCG ACACAACACC CGATCCTGGC CGGGCACCAA	240
CGTGTCGAG AAGAGGTTCT GCACAATTTT GCCAGGGGAC GTAGCACAGT GCTCGAGATA	300
GGGCCGTCTC TGCACAGCGC ACTTAAGCTA CATGGGGCAC CGAACGCCCC CGTCGCAGAC	360
TATCACGGGT GCACCAAGTA CGGCACCCGC GACGGCTCGC GACACATTAC GGCCTTAGAG	420
TCTAGATCCG TCGCCACAGG CCGGCCCCGAG TTCAAGGCCG ACGCCTCACT GCTCGCCAAC	480
GGCATTGCCT CCCGCACCTT CTGCGTCGAC GGAGTCGGCT CTTGCGCGTT CAAATCGCGC	540
GTTGGAATTG CCAATCACTC CCTCTATGAC GTGACCCTAG AGGAGCTGGC CAATGCGTTT	600
GAGAACCACG GACTTCACAT GGTCCGCGCG TTCATGCACA TGCCAGAAGA GCTGCTCTAC	660
ATGGACAACG TGGTTAATGC CGAGCTCGGC TACCGCTTCC ACGTTATTGA AGAGCCTATG	720
GCTGTGAAGG ACTGCGCATT CCAGGGGGGG GACCTCCGTC TCCACTTCCC TGAGTTGGAC	780
TTCATCAACG AGAGCCAAGA GCGGCGCATC GAGAGGCTGG CCGCCCGCGG CTCCTACTCC	840

AGACGCGCCG	TCATTTTCTC	CGGCGACGAC	GACTGGGGTG	ATGCGTACTT	ACACGACTTC	900
CACACATGGC	TCGCCTACCT	ACTGGTGAGG	AACTACCCCA	CTCCGTTTGG	TTTCTCACTC	960
CATATAGAAG	TCCAGAGGCG	CCACGGCTCC	AGCATTGAGC	TGCGCATCAC	TCGCGCGCCA	1020
CCTGGAGACC	GCATGCTGGC	CGTCGTCCCA	AGGACGTCCC	AAGGCCTCTG	CAGAATCCCA	1080
AACATCTTTT	ATTACGCCGA	CGCGTCGGGC	ACTGAGCATA	AGACCATCCT	TACGTCACAG	1140
CACAAAGTCA	ACATGCTGCT	CAATTTTATG	CAAACGCGTC	CTGAGAAGGA	ACTAGTCGAC	1200
ATGACCGTCT	TGATGTCGTT	CGCGCGCGCT	AGGCTGCGCG	CGATCGTGGT	CGCCTCAGAA	1260
GTCACCGAGA	GCTCCTGGAA	CATCTCACCG	GCTGACCTGG	TCCGCACTGT	CGTGTCTCTT	1320
TACGTCTCTC	ACATCATCGA	GCGCCGAAGG	GCTGCGGTCTG	CTGTCAAGAC	CGCCAAGGAC	1380
GACGTCTTTG	GAGAGACTTC	GTTCTGGGAG	AGTCTCAAGC	ACGTCTTGGG	CTCCTGTTGC	1440
GGTCTGCGCA	ACCTCAAAGG	CACCGACGTC	GTCTTTACTA	AGCGCGTCGT	CGATAAGTAC	1500
CGAGTCCACT	CGCTCGGAGA	CATAATCTGC	GACGTCCGCC	TGTCCCCTGA	ACAGGTCGGC	1560
TTCCTGCCGT	CCCGCGTACC	ACCTGCCCCG	GTCTTTCACG	ACAGGGAAGA	GCTTGAGGTC	1620
CTTCGCGAAG	CTGGCTGCTA	CAACGAACGT	CCGGTACCTT	CCACTCCTCC	TGTGGAGGAG	1680
CCCCAAGGTT	TCGACGCCGA	CTTGTGGCAC	GCGACCGCGG	CCTCACTCCC	CGAGTACCGC	1740
GCCACCTTGC	AGGCAGGTCT	CAACACCGAC	GTCAAGCAGC	TCAAGATCAC	CCTCGAGAAC	1800
GCCCTCAAGA	CCATCGACGG	GCTCACCTC	TCCCCAGTCA	GAGGCCTCGA	GATGTACGAG	1860
GGCCCCGCCAG	GCAGCGGCAA	GACGGGCACC	CTCATCGCCG	CCCTTGAGGC	CGCGGGCGGT	1920
AAAGCACTTT	ACGTGGCACC	CACCAGAGAA	CTGAGAGAGG	CTATGGACCG	GCGGATCAAA	1980
CCGCCGTCCG	CCTCGGCTAC	GCAACATGTC	GCCCTTGCGA	TTCTCCGTCG	TGCCACCGCC	2040
GAGGGCGCCC	CTTTCGCTAC	CGTGGTTATC	GACGAGTGCT	TCATGTTCCC	GCTCGTGTAC	2100
GTCGCGATCG	TGCACGCCTT	GTCCCCGAGC	TCACGAATAG	TCCTTGTAGG	GGACGTCCAC	2160
CAAATCGGGT	TTATAGACTT	CCAAGGCACA	AGCGCGAACA	TGCCGCTCGT	TCGCGACGTC	2220
GTTAAGCAGT	GCCGTCGGCG	CACTTTCAAC	CAAACCAAGC	GCTGTCCGGC	CGACGTCGTT	2280
GCCACCACGT	TTTTCCAGAG	CTTGTACCCC	GGGTGCACAA	CCACCTCAGG	GTGCGTCGCA	2340
TCCATCAGCC	ACGTCGCCCC	AGACTACCGC	AACAGCCAGG	CGCAAACGCT	CTGCTTCACG	2400
CAGGAGGAAA	AGTCGCGCCA	CGGGGCTGAG	GGCGCGATGA	CTGTGCACGA	AGCGCAGGGA	2460
CGCACTTTTG	CGTCTGTCAT	TCTGCATTAC	AACGGCTCCA	CAGCAGAGCA	GAAGCTCCTC	2520

GCTGAGAAGT	CGCACCTTCT	AGTCGGCATC	ACGCGCCACA	CCAACCACCT	GTACATCCGC	2580
GACCCGACAG	GTGACATTGA	GAGACAACTC	AACCATAGCG	CGAAAGCCGA	GGTGTTTACA	2640
GACATCCCTG	CACCCCTGGA	GATCACGACT	GTCAAACCGA	GTGAAGAGGT	GCAGCGCAAC	2700
GAAGTGATGG	CAACGATACC	CCCGCAGAGT	GCCACGCCGC	ACGGAGCAAT	CCATCTGCTC	2760
CGCAAGAACT	TCGGGGACCA	ACCCGACTGT	GGCTGTGTCTG	CTTTGGCGAA	GACCGGCTAC	2820
GAGGTGTTTG	GCGGTCTGTC	CAAAATCAAC	GTAAGAGCTTG	CCGAACCCGA	CGCGACCCCG	2880
AAGCCGCATA	GGGCGTTCCA	GGAAGGGGTA	CAGTGGGTCA	AGGTCACCAA	CGCGTCTAAC	2940
AAACACCAGG	CGCTCCAGAC	GCTGTTGTCC	CGCTACACCA	AGCGAAGCGC	TGACCTGCCG	3000
CTACACGAAG	CTAAGGAGGA	CGTCAAACGC	ATGCTAAACT	CGCTTGACCG	ACATTGGGAC	3060
TGGACTGTCA	CTGAAGACGC	CCGTGACCGA	GCTGTCTTCG	AGACCCAGCT	CAAGTTCACC	3120
CAACGCGGCG	GCACCGTCGA	AGACCTGCTG	GAGCCAGACG	ACCCCTACAT	CCGTGACATA	3180
GACTTCCTTA	TGAAGACTCA	GCAGAAAGTG	TCGCCAAGC	CGATCAATAC	GGGCAAGGTC	3240
GGGCAGGGGA	TCGCCGCTCA	CTCAAAGTCT	CTCAACTTCG	TCCTCGCCGC	TTGGATACGC	3300
ATACTCGAGG	AGATACTCCG	TACCGGGAGC	CGCACGGTCC	GGTACAGCAA	CGGTCTCCCC	3360
GACGAAGAAG	AGGCCATGCT	GCTCGAAGCG	AAGATCAATC	AAGTCCCACA	CGCCACGTTC	3420
GTCTCGGCGG	ACTGGACCGA	GTTTGACACC	GCCCACAATA	ACACGAGTGA	GCTGCTCTTC	3480
GCCGCCCTTT	TAGAGCGCAT	CGGCACGCCT	GCAGCTGCCG	TTAATCTATT	CAGAGAACGG	3540
TGTGGGAAAC	GCACCTTGCG	AGCGAAGGGT	CTAGGCTCCG	TTGAAGTCGA	CGGTCTGCTC	3600
GACTCCGGCG	CAGCTTGAC	GCCTTGCCGC	AACACCATCT	TCTCTGCCGC	CGTCATGCTC	3660
ACGCTCTTCC	GCGGCGTCAA	GTTGCGAGCT	TTCAAAGGCG	ACGACTCGCT	CCTCTGTGGT	3720
AGCCATTACC	TCCGTTTCGA	CGCTAGCCGC	CTTCACATGG	GCGAACGTTA	CAAGACCAAA	3780
CATTTGAAGG	TCGAGGTGCA	GAAAAATCGTG	CCGTACATCG	GACTCCTCGT	CTCCGCTGAG	3840
CAGGTCGTCC	TCGACCCTGT	CAGGAGCGCT	CTCAAGATAT	TTGGGCGCTG	CTACACAAGC	3900
GAATCCTTT	ACTCCAAGTA	CGTGGAGGCT	GTGAGAGACA	TCACCAAGGG	CTGGAGTGAC	3960
GCCCGCTACC	ACAGCCTCCT	GTGCCACATG	TCAGCATGCT	ACTACAATTA	CGCGCCGGAG	4020
TCTGCGGCGT	ACATCATCGA	CGCTGTTGTT	CGCTTTGGGC	GCGGCGACTT	CCCGTTTGAA	4080
CAACTGCGCG	TGGTGCGTGC	CCATGTGCAG	GCACCCGACG	CTTACAGCAG	CACGTATCCG	4140
GCTAACGTGC	GCGCATCGTG	CCTTGACCAC	GTCTTCGAGC	CCCGCCAGGC	CGCCGCCCCG	4200

GCAGGTTTCG TTGCGAC ATG TGC GAA GCC GGA AAC GCC TTC TTC ACT TAC	4250
Met Cys Glu Ala Gly Asn Ala Phe Phe Thr Tyr	
1 5 10	
CGC GAA AGC TGG TGT TTC TGC GAC TAC AAG CCA CGT TGC GAC TGG GAC	4298
Arg Glu Ser Trp Cys Phe Cys Asp Tyr Lys Pro Arg Cys Asp Trp Asp	
15 20 25	
TGC GCC CCC GGA GTC TCC ATG GGA TGC ACC TGC AGC CAA CAG CTT TTC	4346
Cys Ala Pro Gly Val Ser Met Gly Cys Thr Cys Ser Gln Gln Leu Phe	
30 35 40	
GGA GTT ATT GAC ACC GGA GAC CCC GTC CAC ATC ATC CTC GCC GTC ATC	4394
Gly Val Ile Asp Thr Gly Asp Pro Val His Ile Ile Leu Ala Val Ile	
45 50 55	
GTC TTC ATC GGA CTC CTC TAC ATC GTG TGG AAG GTC GCT CAG TGG TGG	4442
Val Phe Ile Gly Leu Leu Tyr Ile Val Trp Lys Val Ala Gln Trp Trp	
60 65 70 75	
AGA CAC CGC AAG GAC CAC AGA AGA CTT GAA CAG CAG AAA GCC GCC TTC	4490
Arg His Arg Lys Asp His Arg Arg Leu Glu Gln Gln Lys Ala Ala Phe	
80 85 90	
GCA AGA CAG GCA ATC ACG CTC GTC TGAATGTCTG GACAGAAGCG GAGAAAGGAC	4544
Ala Arg Gln Ala Ile Thr Leu Val	
95	
AGGCAGTTCG TTAAGTCCCC CCACTGCTCC GAGCCCCCTCA TTCTCATTTT CGGAAAGAGC	4604
TCGACTGGCG ACCGGGCGCA CTGTCGCCGC TGCACATCA CCTTCGGCAA CCCCATCCTG	4664
CGCCACGGAC CAGGTTGCCG CGAGGACCAC GCCGGACTTT GCGCCTTTCC TGGGTTCCCA	4724
GTCTGCCCCGT GCTGTCTCGA AGCCGTACCG GCGCCCCACG ACTGCCCCGT GAAAGAAGT	4784
CACCCCGCTC CACGCGTGGA AGGGCGTGAC CGGAGACCGA CCGGAAGTCA GGGAGGACCC	4844
GGAGACAGCG GCGGTCGTCC AGGCTCTGAT CAGCGGCCGT TATCCTCAGA AGACGAAGCT	4904
TTCTCCGAC GCATCCAAAG GCTACTCAAG AACTAAGGGA TGCTCACAAT CCACCTCTTT	4964
TCCTGCCCCG AGTGCGGATT ACCAGGCCCG CGACTGCCAG ACAGTCCGAG TCTGCCGCGC	5024
CGCTGCAGAG ATGGCGCGCT CATGTATTCA CGAGCCGTTG GCTTCATCTG CCGCCAGTGC	5084
CGACTTGAAG CGCATACGCT CTACCTCGGA CTCTGTTCCC GATGTAAAGA TCAGCAAGAG	5144
CGCATGAAGG AACAAAATTA GTTTCCTTGT TCGTAAACAA GGTGGTCCCT CCCATTGAGG	5204
TAAAGACTCT GGTGAGTCCT CAACGTTACT CGTTGAGTCT GCTGCGGTTC GATTCCATTC	5264
CCAAGCAGCA AAGGGTGCGC AACTAGTACG GCGCCCCCTG GGATACCA	5312

(2) INFORMATION FOR SEQ ID NO:42:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 99 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:42:

```

Met Cys Glu Ala Gly Asn Ala Phe Phe Thr Tyr Arg Glu Ser Trp Cys
 1             5             10             15

Phe Cys Asp Tyr Lys Pro Arg Cys Asp Trp Asp Cys Ala Pro Gly Val
          20             25             30

Ser Met Gly Cys Thr Cys Ser Gln Gln Leu Phe Gly Val Ile Asp Thr
          35             40             45

Gly Asp Pro Val His Ile Ile Leu Ala Val Ile Val Phe Ile Gly Leu
 50             55             60

Leu Tyr Ile Val Trp Lys Val Ala Gln Trp Trp Arg His Arg Lys Asp
 65             70             75             80

His Arg Arg Leu Glu Gln Gln Lys Ala Ala Phe Ala Arg Gln Ala Ile
          85             90             95

Thr Leu Val

```

(2) INFORMATION FOR SEQ ID NO:43:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 5312 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: unknown
- (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: DNA

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 4518..4937

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:43:

```

GTTCTGCCTC CCCCGGACGG TAAATATAGG GGAACAATGT ACGCGAAAGC GACAGACGTG      60
GCGCGTGTCT ACGCCGCGGC AGATGTCGCC TACGCGAACG TACTGCAGCA GAGAGCAGTC      120
AAGTTGGACT TCGCCCCGCC ACTGAAGGCA CTAGAAACCC TCCACAGACT GTACTATCCG      180
CTGCGCTTCA AAGGGGGCAC TTTACCCCCG ACACAACACC CGATCCTGGC CGGGCACCAA      240

```

CGTGTCGCAG	AAGAGGTTCT	GCACAATTTT	GCCAGGGGAC	GTCACACAGT	GCTCGAGATA	300
GGGCCGTCTC	TGCACAGCGC	ACTTAAGCTA	CATGGGGCAC	CGAACGCCCC	CGTCGCAGAC	360
TATCACGGGT	GCACCAAGTA	CGGCACCCGC	GACGGCTCGC	GACACATTAC	GGCCTTAGAG	420
TCTAGATCCG	TCGCCACAGG	CCGGCCCCGAG	TTCAAGGCCG	ACGCCTCACT	GCTCGCCAAC	480
GGCATTGCCT	CCCGCACCTT	CTGCGTCGAC	GGAGTCGGCT	CTTGCGCGTT	CAAATCGCGC	540
GTTGGAATTG	CCAATCACTC	CCTCTATGAC	GTGACCCTAG	AGGAGCTGGC	CAATGCGTTT	600
GAGAACCACG	GACTTCACAT	GGTCCGCGCG	TTCATGCACA	TGCCAGAAGA	GCTGCTCTAC	660
ATGGACAACG	TGGTTAATGC	CGAGCTCGGC	TACCGCTTCC	ACGTTATTGA	AGAGCCTATG	720
GCTGTGAAGG	ACTGCGCATT	CCAGGGGGGG	GACCTCCGTC	TCCACTTCCC	TGAGTTGGAC	780
TTCATCAACG	AGAGCCAAGA	GCGGCGCATC	GAGAGGCTGG	CCGCCC GCGG	CTCCTACTCC	840
AGACGCGCCG	TCATTTTCTC	CGGCGACGAC	GACTGGGGTG	ATGCGTACTT	ACACGACTTC	900
CACACATGGC	TCGCCTACCT	ACTGGTGAGG	AACTACCCCA	CTCCGTTTGG	TTTCTCACTC	960
CATATAGAAG	TCCAGAGGCG	CCACGGCTCC	AGCATTGAGC	TGCGCATCAC	TCGCGCGCCA	1020
CCTGGAGACC	GCATGCTGGC	CGTCGTCCCA	AGGACGTCCC	AAGGCCTCTG	CAGAATCCCA	1080
AACATCTTTT	ATTACGCCGA	CGCGTCGGGC	ACTGAGCATA	AGACCATCCT	TACGTCACAG	1140
CACAAAGTCA	ACATGCTGCT	CAATTTTATG	CAAACGCGTC	CTGAGAAGGA	ACTAGTCGAC	1200
ATGACCGTCT	TGATGTCGTT	CGCGCGCGCT	AGGCTGCGCG	CGATCGTGGT	CGCCTCAGAA	1260
GTCACCGAGA	GCTCCTGGAA	CATCTCACCG	GCTGACCTGG	TCCGCACTGT	CGTGTCTCTT	1320
TACGTCTCTC	ACATCATCGA	GCGCCGAAGG	GCTGCGGTCTG	CTGTCAAGAC	CGCCAAGGAC	1380
GACGTCTTTG	GAGAGACTTC	GTTCTGGGAG	AGTCTCAAGC	ACGTCTTGGG	CTCCTGTTGC	1440
GGTCTGCGCA	ACCTCAAAGG	CACCGACGTC	GTCTTTACTA	AGCGCGTCGT	CGATAAGTAC	1500
CGAGTCCACT	CGCTCGGAGA	CATAATCTGC	GACGTCCGCC	TGTCCCCTGA	ACAGGTCGGC	1560
TTCTTGCCGT	CCCGCGTACC	ACCTGCCCCG	GTCTTTACAG	ACAGGGAAGA	GCTTGAGGTC	1620
CTTCGCGAAG	CTGGCTGCTA	CAACGAACGT	CCGGTACCTT	CCACTCCTCC	TGTGGAGGAG	1680
CCCCAAGGTT	TCGACGCCGA	CTTGTGGCAC	GCGACCGCGG	CCTCACTCCC	CGAGTACCGC	1740
GCCACCTTGC	AGGCAGGTCT	CAACACCGAC	GTCAAGCAGC	TCAAGATCAC	CCTCGAGAAC	1800
GCCCTCAAGA	CCATCGACGG	GCTCACCCCTC	TCCCCAGTCA	GAGGCCTCGA	GATGTACGAG	1860
GGCCCCGCCAG	GCAGCGGCAA	GACGGGCACC	CTCATCGCCG	CCCTTGAGGC	CGCGGGCGGT	1920

AAAGCACTTT	ACGTGGCACC	CACCAGAGAA	CTGAGAGAGG	CTATGGACCG	GCGGATCAAA	1980
CCGCCGTCCG	CCTCGGCTAC	GCAACATGTC	GCCCTTGCGA	TTCTCCGTCG	TGCCACCGCC	2040
GAGGGCGCCC	CTTTCGCTAC	CGTGGTTATC	GACGAGTGCT	TCATGTTCCC	GCTCGTGTAC	2100
GTCGCGATCG	TGCACGCCTT	GTCCCCGAGC	TCACGAATAG	TCCTTGTAGG	GGACGTCCAC	2160
CAAATCGGGT	TTATAGACTT	CCAAGGCACA	AGCGCGAACA	TGCCGCTCGT	TCGCGACGTC	2220
GTTAAGCAGT	GCCGTCGGCG	CACTTTCAAC	CAAACCAAGC	GCTGTCCGGC	CGACGTGCTT	2280
GCCACCACGT	TTTTCCAGAG	CTTGTACCCC	GGGTGCACAA	CCACCTCAGG	GTGCGTCGCA	2340
TCCATCAGCC	ACGTCGCCCC	AGACTACCGC	AACAGCCAGG	CGCAAACGCT	CTGCTTCACG	2400
CAGGAGGAAA	AGTCGCGCCA	CGGGGCTGAG	GGCGCGATGA	CTGTGCACGA	AGCGCAGGGA	2460
CGCACTTTTG	CGTCTGTCAT	TCTGCATTAC	AACGGGCTCCA	CAGCAGAGCA	GAAGCTCCTC	2520
GCTGAGAAGT	CGCACCTTCT	AGTCGGCATC	ACGCGCCACA	CCAACCACCT	GTACATCCGC	2580
GACCCGACAG	GTGACATTGA	GAGACAACTC	AACCATAGCG	CGAAAGCCGA	GGTGTTTACA	2640
GACATCCCTG	CACCCCTGGA	GATCACGACT	GTCAAACCGA	GTGAAGAGGT	GCAGCGCAAC	2700
GAAGTGATGG	CAACGATACC	CCCGCAGAGT	GCCACGCCGC	ACGGAGCAAT	CCATCTGCTC	2760
CGCAAGAACT	TCGGGGACCA	ACCCGACTGT	GGCTGTGTCG	CTTTGGCGAA	GACCGGCTAC	2820
GAGGTGTTTG	GCGGTCGTGC	CAAAATCAAC	GTAAGAGCTTG	CCGAACCCGA	CGCGACCCCG	2880
AAGCCGCATA	GGGCGTTCCA	GGAAGGGGTA	CAGTGGGTCA	AGGTCACCAA	CGCGTCTAAC	2940
AAACACCAGG	CGCTCCAGAC	GCTGTTGTCC	CGCTACACCA	AGCGAAGCGC	TGACCTGCCG	3000
CTACACGAAG	CTAAGGAGGA	CGTCAAACGC	ATGCTAAACT	CGCTTGACCG	ACATTGGGAC	3060
TGGACTGTCA	CTGAAGACGC	CCGTGACCGA	GCTGTCTTCG	AGACCCAGCT	CAAGTTCACC	3120
CAACGCGGCG	GCACCGTCGA	AGACCTGCTG	GAGCCAGACG	ACCCCTACAT	CCGTGACATA	3180
GACTTCCTTA	TGAAGACTCA	GCAGAAAAGT	TCGCCCCAAGC	CGATCAATAC	GGGCAAGGTC	3240
GGGCAGGGGA	TCGCCGCTCA	CTCAAAGTCT	CTCAACTTCG	TCCTCGCCGC	TTGGATACGC	3300
ATACTCGAGG	AGATACTCCG	TACCGGGAGC	CGCACGGTCC	GGTACAGCAA	CGGTCTCCCC	3360
GACGAAGAAG	AGGCCATGCT	GCTCGAAGCG	AAGATCAATC	AAGTCCCACA	CGCCACGTTC	3420
GTCTCGGCGG	ACTGGACCGA	GTTTGACACC	GCCCACAATA	ACACGAGTGA	GCTGCTCTTC	3480
GCCGCCCTTT	TAGAGCGCAT	CGGCACGCCT	GCAGCTGCCG	TTAATCTATT	CAGAGAACGG	3540
TGTGGGAAAC	GCACCTTGCG	AGCGAAGGGT	CTAGGCTCCG	TTGAAGTCGA	CGGTCTGCTC	3600

GACTCCGGCG	CAGCTTGGAC	GCCTTGCCGC	AACACCATCT	TCTCTGCCGC	CGTCATGCTC		3660									
ACGCTCTTCC	GCGGCGTCAA	GTTCGCAGCT	TTCAAAGGCG	ACGACTCGCT	CCTCTGTGGT		3720									
AGCCATTACC	TCCGTTTTCGA	CGCTAGCCGC	CTTCACATGG	GCGAACGTTA	CAAGACCAAA		3780									
CATTTGAAGG	TCGAGGTGCA	GAAAATCGTG	CCGTACATCG	GACTCCTCGT	CTCCGCTGAG		3840									
CAGGTCGTCC	TCGACCCTGT	CAGGAGCGCT	CTCAAGATAT	TTGGGCGCTG	CTACACAAGC		3900									
GAACTCCTTT	ACTCCAAGTA	CGTGGAGGCT	GTGAGAGACA	TCACCAAGGG	CTGGAGTGAC		3960									
GCCCGCTACC	ACAGCCTCCT	GTGCCACATG	TCAGCATGCT	ACTACAATTA	CGCGCCGGAG		4020									
TCTGCGGCGT	ACATCATCGA	CGCTGTTGTT	CGCTTTGGGC	GCGGCGACTT	CCCGTTTGAA		4080									
CAACTGCGCG	TGGTGCGTGC	CCATGTGCAG	GCACCCGACG	CTTACAGCAG	CACGTATCCG		4140									
GCTAACGTGC	GCGCATCGTG	CCTTGACCAC	GTCTTCGAGC	CCCGCCAGGC	CGCCGCCCCG		4200									
GCAGGTTTCG	TTGCGACATG	TGCGAAGCCG	GAAACGCCTT	CTTCACTTAC	CGCGAAAGCT		4260									
GGTGTTTCTG	CGACTACAAG	CCACGTTGCG	ACTGGGACTG	CGCCCCGGA	GTCTCCATGG		4320									
GATGCACCTG	CAGCCAACAG	CTTTTCGGAG	TTATTGACAC	CGGAGACCCC	GTCCACATCA		4380									
TCCTCGCCGT	CATCGTCTTC	ATCGGACTCC	TCTACATCGT	GTGGAAGGTC	GCTCAGTGGT		4440									
GGAGACACCG	CAAGGACCAC	AGAAGACTTG	AACAGCAGAA	AGCCGCCTTC	GCAAGACAGG		4500									
CAATCACGCT	CGTCTGA	ATG	TCT	GGA	CAG	AAG	CGG	AGA	AAG	GAC	AGG	CAG		4550		
		Met	Ser	Gly	Gln	Lys	Arg	Arg	Lys	Asp	Arg	Gln				
		1					5					10				
TTC	GTT	AAC	TGC	CCC	CAC	TGC	TCC	GAG	CCC	CTC	ATT	CTC	ATT	TTC	GGA	4598
Phe	Val	Asn	Cys	Pro	His	Cys	Ser	Glu	Pro	Leu	Ile	Leu	Ile	Phe	Gly	
			15					20					25			
AAG	AGC	TCG	ACT	GGC	GAC	CGG	GCC	GAC	TGT	CGC	CGC	TGC	GAC	ATC	ACC	4646
Lys	Ser	Ser	Thr	Gly	Asp	Arg	Ala	Asp	Cys	Arg	Arg	Cys	Asp	Ile	Thr	
			30				35					40				
TTC	GGC	AAC	CCC	ATC	CTG	CGC	CAC	GGA	CCA	GGT	TGC	CGC	GAG	GAC	CAC	4694
Phe	Gly	Asn	Pro	Ile	Leu	Arg	His	Gly	Pro	Gly	Cys	Arg	Glu	Asp	His	
			45				50				55					
GCC	GGA	CTT	TGC	GCC	TTT	CCT	GGG	TTC	CCA	GTC	TGC	CCG	TGC	TGT	CTC	4742
Ala	Gly	Leu	Cys	Ala	Phe	Pro	Gly	Phe	Pro	Val	Cys	Pro	Cys	Cys	Leu	
			60			65				70					75	
GAA	GCC	GTA	CCG	GCC	CCC	CAC	GAC	TGC	CCG	TTG	GAA	AGA	AGT	CAC	CCC	4790
Glu	Ala	Val	Pro	Ala	Pro	His	Asp	Cys	Pro	Leu	Glu	Arg	Ser	His	Pro	
						80				85					90	

GCT CCA CGC GTG GAA GGG CGT GAC CGG AGA CCG ACC GGA AGT CAG GGA	4838
Ala Pro Arg Val Glu Gly Arg Asp Arg Arg Pro Thr Gly Ser Gln Gly	
95 100 105	
GGA CCC GGA GAC AGC GGC GGT CGT CCA GGC TCT GAT CAG CGG CCG TTA	4886
Gly Pro Gly Asp Ser Gly Gly Arg Pro Gly Ser Asp Gln Arg Pro Leu	
110 115 120	
TCC TCA GAA GAC GAA GCT TTC CTC CGA CGC ATC CAA AGG CTA CTC AAG	4934
Ser Ser Glu Asp Glu Ala Phe Leu Arg Arg Ile Gln Arg Leu Leu Lys	
125 130 135	
AAC TAAGGGATGC TCACAATCCA CCTCTTTTCC TGCCCCGAGT GCGGATTACC	4987
Asn	
140	
AGGCCCCGCGA CTGCCAGACA GTCCGAGTCT GCCGCGCCGC TGCAGAGATG GCGCGCTCAT	5047
GTATTCACGA GCCGTTGGCT TCATCTGCCG CCAGTGCCGA CTTGAAGCGC ATACGCTCTA	5107
CCTCGGACTC TGTTCCCGAT GTAAAGATCA GCAAGAGCGC ATGAAGGAAC AAAATTAGTT	5167
TCCTTGTTTCG TAAACAAGGT GGTCCCTCCC ATTGAGGTAA AGACTCTGGT GAGTCCTCAA	5227
CGTTACTCGT TGAGTCTGCT GCGGTTTCGAT TCCATTCCCA AGCAGCAAAG GGTGCGCAAC	5287
TAGTACGGCG CCCCCTGGGA TACCA	5312

(2) INFORMATION FOR SEQ ID NO:44:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 140 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:44:

Met Ser Gly Gln Lys Arg Arg Lys Asp Arg Gln Phe Val Asn Cys Pro	
1 5 10 15	
His Cys Ser Glu Pro Leu Ile Leu Ile Phe Gly Lys Ser Ser Thr Gly	
20 25 30	
Asp Arg Ala Asp Cys Arg Arg Cys Asp Ile Thr Phe Gly Asn Pro Ile	
35 40 45	
Leu Arg His Gly Pro Gly Cys Arg Glu Asp His Ala Gly Leu Cys Ala	
50 55 60	
Phe Pro Gly Phe Pro Val Cys Pro Cys Cys Leu Glu Ala Val Pro Ala	
65 70 75 80	

Pro His Asp Cys Pro Leu Glu Arg Ser His Pro Ala Pro Arg Val Glu
 85 90 95

Gly Arg Asp Arg Arg Pro Thr Gly Ser Gln Gly Gly Pro Gly Asp Ser
 100 105 110

Gly Gly Arg Pro Gly Ser Asp Gln Arg Pro Leu Ser Ser Glu Asp Glu
 115 120 125

Ala Phe Leu Arg Arg Ile Gln Arg Leu Leu Lys Asn
 130 135 140

(2) INFORMATION FOR SEQ ID NO:45:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 5368 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: unknown
- (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: DNA

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 4944..5162

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:45:

GTTCTGCCTC CCCCGACGG TAAATATAGG GGAACAATGT ACGCGAAAGC GACAGACGTG	60
GCGCGTGTCT ACGCCGCGGC AGATGTCGCC TACGCGAACG TACTGCAGCA GAGAGCAGTC	120
AAGTTGGA CT TCGCCCCGCC ACTGAAGGCA CTAGAAACCC TCCACAGACT GTACTATCCG	180
CTGCGCTTCA AAGGGGGCAC TTTACCCCCG ACACAACACC CGATCCTGGC CGGGCACCAA	240
CGTGTCGCAG AAGAGGTTCT GCACAATTTT GCCAGGGGAC GTAGCACAGT GCTCGAGATA	300
GGGCCGTCTC TGCACAGCGC ACTTAAGCTA CATGGGGCAC CGAACGCCCC CGTCGCAGAC	360
TATCACGGGT GCACCAAGTA CGGCACCCGC GACGGCTCGC GACACATTAC GGCCTTAGAG	420
TCTAGATCCG TCGCCACAGG CCGGCCCCGAG TTCAAGGCCG ACGCCTCACT GCTCGCCAAC	480
GGCATTCGCT CCCGCACCTT CTGCGTCGAC GGAGTCGGCT CTTGCGCGTT CAAATCGCGC	540
GTTGGAATTG CCAATCACTC CCTCTATGAC GTGACCCTAG AGGAGCTGGC CAATGCGTTT	600
GAGAACCACG GACTTCACAT GGTCCGCGCG TTCATGCACA TGCCAGAAGA GCTGCTCTAC	660
ATGGACAACG TGGTTAATGC CGAGCTCGGC TACCGCTTCC ACGTTATTGA AGAGCCTATG	720
GCTGTGAAGG ACTGCGCATT CCAGGGGGGG GACCTCCGTC TCCACTTCCC TGAGTTGGAG	780
TTCATCAACG AGAGCCAAGA GCGGCGCATC GAGAGGCTGG CCGCCCGCGG CTCCTACTCC	840

AGACGCGCCG	TCATTTTCTC	CGGCGACGAC	GACTGGGGTG	ATGCGTACTT	ACACGACTTC	900
CACACATGGC	TCGCCTACCT	ACTGGTGAGG	AACTACCCCA	CTCCGTTTGG	TTTCTCACTC	960
CATATAGAAG	TCCAGAGGCG	CCACGGCTCC	AGCATTGAGC	TGCGCATCAC	TCGCGCGCCA	1020
CCTGGAGACC	GCATGCTGGC	CGTCGTCCCA	AGGACGTCCC	AAGGCCTCTG	CAGAATCCCA	1080
AACATCTTTT	ATTACGCCGA	CGCGTCGGGC	ACTGAGCATA	AGACCATCCT	TACGTCACAG	1140
CACAAAGTCA	ACATGCTGCT	CAATTTTATG	CAAACGCGTC	CTGAGAAGGA	ACTAGTCGAC	1200
ATGACCGTCT	TGATGTCGTT	CGCGCGCGCT	AGGCTGCGCG	CGATCGTGGT	CGCCTCAGAA	1260
GTCACCGAGA	GCTCCTGGAA	CATCTCACCG	GCTGACCTGG	TCCGCACTGT	CGTGTCTCTT	1320
TACGTCCTCC	ACATCATCGA	GCGCCGAAGG	GCTGCGGTCTG	CTGTCAAGAC	CGCCAAGGAC	1380
GACGTCTTTG	GAGAGACTTC	GTTCTGGGAG	AGTCTCAAGC	ACGTCTTGGG	CTCCTGTTGC	1440
GGTCTGCGCA	ACCTCAAAGG	CACCGACGTC	GTCTTTACTA	AGCGCGTCGT	CGATAAGTAC	1500
CGAGTCCACT	CGCTCGGAGA	CATAATCTGC	GACGTCCGCC	TGTCCCCTGA	ACAGGTCGGC	1560
TTCCTGCCGT	CCCGCGTACC	ACCTGCCCCG	GTCTTTCACG	ACAGGGAAGA	GCTTGAGGTC	1620
CTTCGCGAAG	CTGGCTGCTA	CAACGAACGT	CCGGTACCTT	CCACTCCTCC	TGTGGAGGAG	1680
CCCCAAGGTT	TCGACGCCGA	CTTGTGGCAC	GCGACCGCGG	CCTCACTCCC	CGAGTACCGC	1740
GCCACCTTGC	AGGCAGGTCT	CAACACCGAC	GTCAAGCAGC	TCAAGATCAC	CCTCGAGAAC	1800
GCCCTCAAGA	CCATCGACGG	GCTCACCTC	TCCCCAGTCA	GAGGCCTCGA	GATGTACGAG	1860
GGCCCGCCAG	GCAGCGGCAA	GACGGGCACC	CTCATCGCCG	CCCTTGAGGC	CGCGGGCGGT	1920
AAAGCACTTT	ACGTGGCACC	CACCAGAGAA	CTGAGAGAGG	CTATGGACCG	GCGGATCAAA	1980
CCGCCGTCCG	CCTCGGCTAC	GCAACATGTC	GCCCTTGCGA	TTCTCCGTCG	TGCCACCGCC	2040
GAGGGCGCCC	CTTTCGCTAC	CGTGTTTATC	GACGAGTGCT	TCATGTTCCC	GCTCGTGTAC	2100
GTCGCGATCG	TGCACGCCTT	GTCCCCGAGC	TCACGAATAG	TCCTTGTAGG	GGACGTCCAC	2160
CAAATCGGGT	TTATAGACTT	CCAAGGCACA	AGCGCGAACA	TGCCGCTCGT	TCGCGACGTC	2220
GTTAAGCAGT	GCCGTCGGCG	CACTTTCAAC	CAAACCAAGC	GCTGTCCGGC	CGACGTGCTT	2280
GCCACCACGT	TTTTCCAGAG	CTTGTACCCC	GGGTGCACAA	CCACCTCAGG	GTGCGTCGCA	2340
TCCATCAGCC	ACGTCGCCCC	AGACTACCGC	AACAGCCAGG	CGCAAACGCT	CTGCTTCACG	2400
CAGGAGGAAA	AGTCGCGCCA	CGGGGCTGAG	GGCGCGATGA	CTGTGCACGA	AGCGCAGGGA	2460
CGCACTTTTG	CGTCTGTCAT	TCTGCATTAC	AACGGCTCCA	CAGCAGAGCA	GAAGCTCCTC	2520

GCTGAGAAGT	CGCACCTTCT	AGTCGGCATC	ACGCGCCACA	CCAACCACCT	GTACATCCGC	2580
GACCCGACAG	GTGACATTGA	GAGACAACTC	AACCATAGCG	CGAAAGCCGA	GGTGTTTACA	2640
GACATCCCTG	CACCCCTGGA	GATCACGACT	GTCAAACCGA	GTGAAGAGGT	GCAGCGCAAC	2700
GAAGTGATGG	CAACGATACC	CCCGCAGAGT	GCCACGCCGC	ACGGAGCAAT	CCATCTGCTC	2760
CGCAAGAACT	TCGGGGACCA	ACCCGACTGT	GGCTGTGTCTG	CTTTGGCGAA	GACCGGCTAC	2820
GAGGTGTTTTG	GCGGTCGTGC	CAAAATCAAC	GTAAGAGCTTG	CCGAACCCGA	CGCGACCCCG	2880
AAGCCGCATA	GGGCGTTCCA	GGAAGGGGTA	CAGTGGGTCA	AGGTCACCAA	CGCGTCTAAC	2940
AAACACCAGG	CGCTCCAGAC	GCTGTTGTCC	CGCTACACCA	AGCGAAGCGC	TGACCTGCCG	3000
CTACACGAAG	CTAAGGAGGA	CGTCAAACGC	ATGCTAAACT	CGCTTGACCG	ACATTGGGAC	3060
TGGACTGTCA	CTGAAGACGC	CCGTGACCGA	GCTGTCTTTCG	AGACCCAGCT	CAAGTTCACC	3120
CAACGCGGCG	GCACCGTCGA	AGACCTGCTG	GAGCCAGACG	ACCCCTACAT	CCGTGACATA	3180
GACTTCCTTA	TGAAGACTCA	GCAGAAAGTG	TCGCCCAAGC	CGATCAATAC	GGGCAAGGTC	3240
GGGCAGGGGA	TCGCCGCTCA	CTCAAAGTCT	CTCAACTTCG	TCCTCGCCGC	TTGGATACGC	3300
ATACTCGAGG	AGATACTCCG	TACCGGGAGC	CGCACGGTCC	GGTACAGCAA	CGGTCTCCCC	3360
GACGAAGAAG	AGGCCATGCT	GCTCGAAGCG	AAGATCAATC	AAGTCCCACA	CGCCACGTTC	3420
GTCTCGGCGG	ACTGGACCGA	GTTTGACACC	GCCCACAATA	ACACGAGTGA	GCTGCTCTTC	3480
GCCGCCCTTT	TAGAGCGCAT	CGGCACGCCT	GCAGCTGCCG	TTAATCTATT	CAGAGAACGG	3540
TGTGGGAAAC	GCACCTTGCG	AGCGAAGGGT	CTAGGCTCCG	TTGAAGTCGA	CGGTCTGCTC	3600
GACTCCGGCG	CAGCTTGAC	GCCTTGCCGC	AACACCATCT	TCTCTGCCGC	CGTCATGCTC	3660
ACGCTCTTCC	GCGGCGTCAA	GTTTCGAGCT	TTCAAAGGCG	ACGACTCGCT	CCTCTGTGGT	3720
AGCCATTACC	TCCGTTTCGA	CGCTAGCCGC	CTTCACATGG	GCGAACGTTA	CAAGACCAAA	3780
CATTTGAAGG	TCGAGGTGCA	GAAAATCGTG	CCGTACATCG	GACTCCTCGT	CTCCGCTGAG	3840
CAGGTCGTCC	TCGACCCTGT	CAGGAGCGCT	CTCAAGATAT	TTGGGCGCTG	CTACACAAGC	3900
GAACTCCTTT	ACTCCAAGTA	CGTGGAGGCT	GTGAGAGACA	TCACCAAGGG	CTGGAGTGAC	3960
GCCCCGTACC	ACAGCCTCCT	GTGCCACATG	TCAGCATGCT	ACTACAATTA	CGCGCCGGAG	4020
TCTGCGGCGT	ACATCATCGA	CGCTGTTGTT	CGCTTTGGGC	GCGGCGACTT	CCCGTTTGAA	4080
CAACTGCGCG	TGGTGCGTGC	CCATGTGCAG	GCACCCGACG	CTTACAGCAG	CACGTATCCG	4140
GCTAACGTGC	GCGCATCGTG	CCTTGACCAC	GTCTTCGAGC	CCCGCCAGGC	CGCCGCCCCG	4200

GCAGGTTTCG TTGCGACATG TCGGAAGCCG GAAACGCCTT CTTCACTTAC CGCGAAAGCT	4260
GGTGTTCCTG CGACTACAAG CCACGTTGCG ACTGGGACTG CGCCCCCGGA GTCTCCATGG	4320
GATGCACCTG CAGCCAACAG CTTTTCGGAG TTATTGACAC CGGAGACCCC GTCCACATCA	4380
TCCTCGCCGT CATCGTCTTC ATCGGACTCC TCTACATCGT GTGGAAGGTC GCTCAGTGGT	4440
GGAGACACCG CAAGGACCAC AGAAGACTTG AACAGCAGAA AGCCGCCTTC GCAAGACAGG	4500
CAATCACGCT CGTCTGAATG TCTGGACAGA AGCGGAGAAA GGACAGGCAG TTCGTAACT	4560
GCCCCACTG CTCCGAGCCC CTCATTCTCA TTTTCGGAAA GAGCTCGACT GGCGACCGGG	4620
CCGACTGTCG CCGCTGCGAC ATCACCTTCG GCAACCCCAT CCTGCGCCAC GGACCAGGTT	4680
GCCGCGAGGA CCACGCCGGA CTTTGCGCCT TTCCTGGGTT CCCAGTCTGC CCGTGCTGTC	4740
TCGAAGCCGT ACCGGCCCCC CACGACTGCC CGTTGGAAAG AAGTCACCCC GCTCCACGCG	4800
TGGAAGGGCG TGACCGGAGA CCGACCGGAA GTCAGGGAGG ACCCGGAGAC AGCGGCGGTC	4860
GTCCAGGCTC TGATCAGCGG CCGTTATCCT CAGAAGACGA AGCTTTCCTC CGACGCATCC	4920
AAAGGCTACT CAAGAACTAA GGG ATG CTC ACA ATC CAC CTC TTT TCC TGC	4970
Met Leu Thr Ile His Leu Phe Ser Cys	
1 5	
TTC CCC GAG TGC GGA TTA CCA GGC CCG CGA CTG CCA GAC AGT CCG AGT	5018
Phe Pro Glu Cys Gly Leu Pro Gly Pro Arg Leu Pro Asp Ser Pro Ser	
10 15 20 25	
CTG GCG GAG ACC GCG CCG CTG CAG AGA TGG CGC GCT CAT GTA TTC ACG	5066
Leu Ala Glu Thr Ala Pro Leu Gln Arg Trp Arg Ala His Val Phe Thr	
30 35 40	
AGC CGT TGG CAG AGC AGA GAA TTA GAA AGT TCA TCT GCC GCC AGT GCC	5114
Ser Arg Trp Gln Ser Arg Glu Leu Glu Ser Ser Ser Ala Ala Ser Ala	
45 50 55	
GAC TTG AAG CGC ATA CGC TCT ACC TCG GAC AGG CAG GAA TTG CTC TGT	5162
Asp Leu Lys Arg Ile Arg Ser Thr Ser Asp Arg Gln Glu Leu Leu Cys	
60 65 70	
TCCCGATGTA AAGATCAGCA AGAGCGCATG AAGGAACAAA ATCAGCAGGG AGTGGATAGT	5222
TTCCTTGTTT GTAAACAAGG TGGTCCCTCC CATTGAGGTA AAGACTCTGG TGAGTCCTCA	5282
ACGTTACTCG TTGAGTCTGC TCGGTTCTGA TTCCATTCCC AAGCAGCAAA GGGTGCGCAA	5342
CTAGTACGGC GCCCCCTGGG ATACCA	5368

(2) INFORMATION FOR SEQ ID NO:46:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 73 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:46:

```

Met Leu Thr Ile His Leu Phe Ser Cys Phe Pro Glu Cys Gly Leu Pro
 1             5             10             15
Gly Pro Arg Leu Pro Asp Ser Pro Ser Leu Ala Glu Thr Ala Pro Leu
                20             25             30
Gln Arg Trp Arg Ala His Val Phe Thr Ser Arg Trp Gln Ser Arg Glu
                35             40             45
Leu Glu Ser Ser Ser Ala Ala Ser Ala Asp Leu Lys Arg Ile Arg Ser
                50             55             60
Thr Ser Asp Arg Gln Glu Leu Leu Cys
 65             70

```

(2) INFORMATION FOR SEQ ID NO:47:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2478 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: unknown
- (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: DNA

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 283..753

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:47:

```

GTTTTTCTTT CTTTACCAAG TGTGGTAAAA TTAAACAAA GAAGAAAACC AGGACCGTAA      60
CCCGGCCCTT ACACACCTCG AGTCCGTGAC CACCGGATTA TACGTCGCCC ACCACACGGC      120
GCCTTTTCCG ACCACTCTCG AGAGTCGTTG GGAGTTTCGT CCGTGACCAC CCGGTTGGCA      180
GTCGACAGAC GCTTCCGGAC CACTAGAACC TCCTCGAGCG ACGCACACAC AGCACACACA      240
CCGCCTTAGC TGCACCTACG GCAGCGTTGA TAGCGCGGAT TT ATG AGC GAG CAC      294
                               Met Ser Glu His
                               1
ACC ATC GCC CAC TCC ATC ACA TTA CCA CCC GGT TAC ACC CTT GCC CTA      342

```

Thr	Ile	Ala	His	Ser	Ile	Thr	Leu	Pro	Pro	Gly	Tyr	Thr	Leu	Ala	Leu	
5					10					15					20	
ATA	CCC	CCT	GAA	CCT	GAA	GCA	GGA	TGG	GAG	ATG	CTG	GAG	TGG	CGT	CAC	390
Ile	Pro	Pro	Glu	Pro	Glu	Ala	Gly	Trp	Glu	Met	Leu	Glu	Trp	Arg	His	
			25					30						35		
AGC	GAC	CTC	ACA	ACC	GTC	GCG	GAA	CCC	GTA	ACG	TTC	GGG	TCA	GCG	CCA	438
Ser	Asp	Leu	Thr	Thr	Val	Ala	Glu	Pro	Val	Thr	Phe	Gly	Ser	Ala	Pro	
			40					45					50			
ACA	CCG	TCA	CCG	TCA	ATG	GTA	GAA	GAA	ACC	AAC	GGC	GTC	GGA	CCG	GAA	486
Thr	Pro	Ser	Pro	Ser	Met	Val	Glu	Glu	Thr	Asn	Gly	Val	Gly	Pro	Glu	
			55				60					65				
GGC	AAG	TTT	CTC	CCC	CTG	ACA	ATT	TCA	CCG	CTG	CTG	CAC	AAG	ACC	TCG	534
Gly	Lys	Phe	Leu	Pro	Leu	Thr	Ile	Ser	Pro	Leu	Leu	His	Lys	Thr	Ser	
	70					75					80					
CGC	AAA	GCC	TTG	ACG	CCA	ACA	CCG	TCA	CTT	TCC	CCG	CTA	ACA	TCT	CTA	582
Arg	Lys	Ala	Leu	Thr	Pro	Thr	Pro	Ser	Leu	Ser	Pro	Leu	Thr	Ser	Leu	
	85				90				95						100	
GCA	TGC	CCG	AAT	TCC	GGA	ATT	GGG	CCA	AGG	GAA	AGA	TCG	ACC	TCG	ACT	630
Ala	Cys	Pro	Asn	Ser	Gly	Ile	Gly	Pro	Arg	Glu	Arg	Ser	Thr	Ser	Thr	
			105					110					115			
CCG	ATT	CCA	TCG	GCT	GGT	ACT	TCA	AGT	ACC	TTG	ACC	CAG	CGG	GTG	CTA	678
Pro	Ile	Pro	Ser	Ala	Gly	Thr	Ser	Ser	Thr	Leu	Thr	Gln	Arg	Val	Leu	
			120					125				130				
CAG	AGT	CTG	CGC	GCG	CCG	TCG	GCG	AGT	ACT	CGA	AGA	TCC	CTG	ACG	GCC	726
Gln	Ser	Leu	Arg	Ala	Pro	Ser	Ala	Ser	Thr	Arg	Arg	Ser	Leu	Thr	Ala	
		135					140					145				
TCG	TCA	AGT	TCT	CCG	TCG	ACG	CAG	AGA	TAAGAGAGAT	CTATAACGAG						773
Ser	Ser	Ser	Ser	Pro	Ser	Thr	Gln	Arg								
	150					155										
GAGT	GCCCCG	TCGT	CACTGA	CGTG	TCCGTC	CCCCT	CGACG	GCCG	CCAGTG	GAGC	CTCTCG					833
ATTTT	CTCCT	TTCC	GATGTT	CAGA	ACCGCC	TACG	TCGCCG	TAGC	GAACGT	CGAG	AACAAG					893
GAGAT	GTTCG	TCGAC	GTTGT	CAAC	GACCTC	ATCG	AGTGGC	TCAACA	ATCT	CGCC	GACTGG					953
CGTTAT	GTTCG	TTGACT	CTGA	ACAG	TGGATT	AACTT	CACCA	ATGAC	ACCAC	GTA	CTACGTC					1013
CGCAT	CCGCG	TTCT	ACGTCC	AACCT	ACGAC	GTTCC	AGACC	CCAC	AGAGGG	CCTT	GTTTCG					1073
ACAGT	CTCAG	ACTAC	CGCCT	CACTT	TATAAG	GCGATA	AACAT	GTGA	AGCCAA	CATG	CCAACA					1133
CTCGT	CGACC	AAGG	CTTTTG	GATCG	GCGGC	CAGTAC	GCTC	TCAC	CCGAC	TAGC	CTACCG					1193
CAGTAC	GACG	TCAG	CGAGGC	CTAC	GCTCTG	CACACT	TTTGA	CCTT	CGCCAG	ACCAT	CCAGC					1253
GCCG	CTGCAC	TCGCG	TTTGT	GTGGG	CAGGT	TTGCC	CACAGG	GTGG	CACTGC	GCCT	GCAGGC					1313

ACTCCAGCCT GGGAGCAGGC ATCCTCGGGT GGCTACCTCA CCTGGCGCCA CAACGGTACT	1373
ACTTTCCAG CTGGCTCCGT TAGCTACGTT CTCCCTGAGG GTTTCGCCCT TGAGCGCTAC	1433
GACCCGAACG ACGGCTCTTG GACCGACTTC GCTTCCGCAG GAGACACCGT CACTTTCCGG	1493
CAGGTCGCCG TCGACGAGGT CGTTGTGACC AACAAACCCCG CCGGCGGCGG CAGCGCCCCC	1553
ACCTTCACCG TGAGAGTGCC CCCTTCAAAC GCTTACACCA ACACCGTGTT TAGGAACACG	1613
CTCTTAGAGA CTCGACCCTC CTCTCGTAGG CTCGAACTCC CTATGCCACC TGCTGACTTT	1673
GGACAGACGG TCGCCAACAA CCCGAAGATC GAGCAGTCGC TTCTTAAAGA AACACTTGGC	1733
TGCTATTTGG TCCACTCCAA AATGCGAAAC CCCGTTTTTC AGCTCACGCC AGCCAGCTCC	1793
TTTGGCGCCG TTTCTTCAA CAATCCGGGT TATGAGCGCA CACGCGACCT CCCGGACTAC	1853
ACTGGCATCC GTGACTCATT CGACCAGAAC ATGTCCACCG CTGTGGCCCA CTTCCGCTCA	1913
CTCTCCCACT CCTGCAGTAT CGTCACTAAG ACCTACCAGG GTTGGGAAGG CGTCACGAAC	1973
GTCAACACGC CTTTCGGCCA ATTCGCGCAC GCGGGCCTCC TCAAGAATGA GGAGATCCTC	2033
TGCCTCGCCG ACGACCTGGC CACCCGTCTC ACAGGTGTCT ACCCCGCCAC TGACAACTTC	2093
GCGGCCGCCG TTTCTGCCTT CGCCGCGAAC ATGCTGTCCT CCGTGCTGAA GTCGGAGGCA	2153
ACGTCCTCCA TCATCAAGTC CGTTGGCGAG ACTGCCGTCG GCGCGGCTCA GTCCGGCCTC	2213
GCGAAGCTAC CCGGACTGCT AATGAGTGTA CCAGGGAAGA TTGCCGCGCG TGTCCGCGCG	2273
CGCCGAGCGC GCCGCCGCGC CGCTCGTGCC AATTAGTTTG CTCGCTCCTG TTTCGCCGTT	2333
TCGTAAAACG GCGTGGTCCC GCACATTACG CGTACCCTAA AGACTCTGGT GAGTCCCCGT	2393
CGTTACACGA CGGGTCTGCC GCGGTTCGAT TCCATTCCCA AGCGGCAAGA AGGACGTAGT	2453
TAGCTCTGCG TCCCTCGGGA TACCA	2478

(2) INFORMATION FOR SEQ ID NO:48:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 157 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:48:

Met	Ser	Glu	His	Thr	Ile	Ala	His	Ser	Ile	Thr	Leu	Pro	Pro	Gly	Tyr
1					5				10					15	

Thr Leu Ala Leu Ile Pro Pro Glu Pro Glu Ala Gly Trp Glu Met Leu
 20 25 30
 Glu Trp Arg His Ser Asp Leu Thr Thr Val Ala Glu Pro Val Thr Phe
 35 40 45
 Gly Ser Ala Pro Thr Pro Ser Pro Ser Met Val Glu Glu Thr Asn Gly
 50 55 60
 Val Gly Pro Glu Gly Lys Phe Leu Pro Leu Thr Ile Ser Pro Leu Leu
 65 70 75 80
 His Lys Thr Ser Arg Lys Ala Leu Thr Pro Thr Pro Ser Leu Ser Pro
 85 90 95
 Leu Thr Ser Leu Ala Cys Pro Asn Ser Gly Ile Gly Pro Arg Glu Arg
 100 105 110
 Ser Thr Ser Thr Pro Ile Pro Ser Ala Gly Thr Ser Ser Thr Leu Thr
 115 120 125
 Gln Arg Val Leu Gln Ser Leu Arg Ala Pro Ser Ala Ser Thr Arg Arg
 130 135 140
 Ser Leu Thr Ala Ser Ser Ser Pro Ser Thr Gln Arg
 145 150 155

(2) INFORMATION FOR SEQ ID NO:49:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2478 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: unknown
- (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: DNA

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 366..2306

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:49:

GTTTTTCTTT CTTTACCAAG TGTGGTAAAA TTAAACAAA GAAGAAAACC AGGACCGTAA 60
 CCCGGCCCTT ACACACCTCG AGTCCGTGAC CACCGGATTA TACGTCGCCC ACCACACGGC 120
 GCCTTTTCCG ACCACTCTCG AGAGTCGTTG GGAGTTTCGT CCGTGACCAC CCGGTTGGCA 180
 GTCGACAGAC GCTTCCGGAC CACTAGAACC TCCTCGAGCG ACGCACACAC AGCACACACA 240
 CCGCCTTAGC TGCACCTACG GCAGCGTTGA TAGCGCGGAT TTATGAGCGA GCACACCATC 300
 GCCCACTCCA TCACATTACC ACCCGGTTAC ACCCTTGCCC TAATACCCCC TGAACCTGAA 360
 GCAGG ATG GGA GAT GCT GGA GTG GCG TCA CAG CGA CCT CAC AAC CGT 407

Met Gly Asp Ala Gly Val Ala Ser Gln Arg Pro His Asn Arg																	
1			5			10											
CGC	GGA	ACC	CGT	AAC	GTT	CGG	GTC	AGC	GCC	AAC	ACC	GTC	ACC	GTC	AAT		455
Arg	Gly	Thr	Arg	Asn	Val	Arg	Val	Ser	Ala	Asn	Thr	Val	Thr	Val	Asn		
15					20					25					30		
GGT	AGA	AGA	AAC	CAA	CGG	CGT	CGG	ACC	GGA	AGG	CAA	GTT	TCT	CCC	CCT		503
Gly	Arg	Arg	Asn	Gln	Arg	Arg	Arg	Thr	Gly	Arg	Gln	Val	Ser	Pro	Pro		
			35						40					45			
GAC	AAT	TTC	ACC	GCT	GCT	GCA	CAA	GAC	CTC	GCG	CAA	AGC	CTT	GAC	GCC		551
Asp	Asn	Phe	Thr	Ala	Ala	Ala	Gln	Asp	Leu	Ala	Gln	Ser	Leu	Asp	Ala		
			50					55					60				
AAC	ACC	GTC	ACT	TTC	CCC	GCT	AAC	ATC	TCT	AGC	ATG	CCC	GAA	TTC	CGG		599
Asn	Thr	Val	Thr	Phe	Pro	Ala	Asn	Ile	Ser	Ser	Met	Pro	Glu	Phe	Arg		
		65					70					75					
AAT	TGG	GCC	AAG	GGA	AAG	ATC	GAC	CTC	GAC	TCC	GAT	TCC	ATC	GGC	TGG		647
Asn	Trp	Ala	Lys	Gly	Lys	Ile	Asp	Leu	Asp	Ser	Asp	Ser	Ile	Gly	Trp		
	80					85					90						
TAC	TTC	AAG	TAC	CTT	GAC	CCA	GCG	GGT	GCT	ACA	GAG	TCT	GCG	CGC	GCC		695
Tyr	Phe	Lys	Tyr	Leu	Asp	Pro	Ala	Gly	Ala	Thr	Glu	Ser	Ala	Arg	Ala		
	95				100					105					110		
GTC	GGC	GAG	TAC	TCG	AAG	ATC	CCT	GAC	GGC	CTC	GTC	AAG	TTC	TCC	GTC		743
Val	Gly	Glu	Tyr	Ser	Lys	Ile	Pro	Asp	Gly	Leu	Val	Lys	Phe	Ser	Val		
				115					120					125			
GAC	GCA	GAG	ATA	AGA	GAG	ATC	TAT	AAC	GAG	GAG	TGC	CCC	GTC	GTC	ACT		791
Asp	Ala	Glu	Ile	Arg	Glu	Ile	Tyr	Asn	Glu	Glu	Cys	Pro	Val	Val	Thr		
			130					135					140				
GAC	GTG	TCC	GTC	CCC	CTC	GAC	GGC	CGC	CAG	TGG	AGC	CTC	TCG	ATT	TTC		839
Asp	Val	Ser	Val	Pro	Leu	Asp	Gly	Arg	Gln	Trp	Ser	Leu	Ser	Ile	Phe		
		145					150					155					
TCC	TTT	CCG	ATG	TTC	AGA	ACC	GCC	TAC	GTC	GCC	GTA	GCG	AAC	GTC	GAG		887
Ser	Phe	Pro	Met	Phe	Arg	Thr	Ala	Tyr	Val	Ala	Val	Ala	Asn	Val	Glu		
	160					165					170						
AAC	AAG	GAG	ATG	TCG	CTC	GAC	GTT	GTC	AAC	GAC	CTC	ATC	GAG	TGG	CTC		935
Asn	Lys	Glu	Met	Ser	Leu	Asp	Val	Val	Asn	Asp	Leu	Ile	Glu	Trp	Leu		
	175				180					185					190		
AAC	AAT	CTC	GCC	GAC	TGG	CGT	TAT	GTC	GTT	GAC	TCT	GAA	CAG	TGG	ATT		983
Asn	Asn	Leu	Ala	Asp	Trp	Arg	Tyr	Val	Val	Asp	Ser	Glu	Gln	Trp	Ile		
				195				200						205			
AAC	TTC	ACC	AAT	GAC	ACC	ACG	TAC	TAC	GTC	CGC	ATC	CGC	GTT	CTA	CGT		1031
Asn	Phe	Thr	Asn	Asp	Thr	Thr	Tyr	Tyr	Val	Arg	Ile	Arg	Val	Leu	Arg		
			210					215					220				
CCA	ACC	TAC	GAC	GTT	CCA	GAC	CCC	ACA	GAG	GGC	CTT	GTT	CGC	ACA	GTC		1079

Pro	Thr	Tyr	Asp	Val	Pro	Asp	Pro	Thr	Glu	Gly	Leu	Val	Arg	Thr	Val		
		225					230					235					
TCA	GAC	TAC	CGC	CTC	ACT	TAT	AAG	GCG	ATA	ACA	TGT	GAA	GCC	AAC	ATG	1127	
Ser	Asp	Tyr	Arg	Leu	Thr	Tyr	Lys	Ala	Ile	Thr	Cys	Glu	Ala	Asn	Met		
		240					245					250					
CCA	ACA	CTC	GTC	GAC	CAA	GGC	TTT	TGG	ATC	GGC	GGC	CAG	TAC	GCT	CTC	1175	
Pro	Thr	Leu	Val	Asp	Gln	Gly	Phe	Trp	Ile	Gly	Gly	Gln	Tyr	Ala	Leu		
		255				260					265				270		
ACC	CCG	ACT	AGC	CTA	CCG	CAG	TAC	GAC	GTC	AGC	GAG	GCC	TAC	GCT	CTG	1223	
Thr	Pro	Thr	Ser	Leu	Pro	Gln	Tyr	Asp	Val	Ser	Glu	Ala	Tyr	Ala	Leu		
				275					280						285		
CAC	ACT	TTG	ACC	TTC	GCC	AGA	CCA	TCC	AGC	GCC	GCT	GCA	CTC	GCG	TTT	1271	
His	Thr	Leu	Thr	Phe	Ala	Arg	Pro	Ser	Ser	Ala	Ala	Ala	Leu	Ala	Phe		
			290						295					300			
GTG	TGG	GCA	GGT	TTG	CCA	CAG	GGT	GGC	ACT	GCG	CCT	GCA	GGC	ACT	CCA	1319	
Val	Trp	Ala	Gly	Leu	Pro	Gln	Gly	Gly	Thr	Ala	Pro	Ala	Gly	Thr	Pro		
		305						310					315				
GCC	TGG	GAG	CAG	GCA	TCC	TCG	GGT	GGC	TAC	CTC	ACC	TGG	CGC	CAC	AAC	1367	
Ala	Trp	Glu	Gln	Ala	Ser	Ser	Gly	Gly	Tyr	Leu	Thr	Trp	Arg	His	Asn		
		320					325					330					
GGT	ACT	ACT	TTC	CCA	GCT	GGC	TCC	GTT	AGC	TAC	GTT	CTC	CCT	GAG	GGT	1415	
Gly	Thr	Thr	Phe	Pro	Ala	Gly	Ser	Val	Ser	Tyr	Val	Leu	Pro	Glu	Gly		
		335				340					345				350		
TTC	GCC	CTT	GAG	CGC	TAC	GAC	CCG	AAC	GAC	GGC	TCT	TGG	ACC	GAC	TTC	1463	
Phe	Ala	Leu	Glu	Arg	Tyr	Asp	Pro	Asn	Asp	Gly	Ser	Trp	Thr	Asp	Phe		
				355						360					365		
GCT	TCC	GCA	GGA	GAC	ACC	GTC	ACT	TTC	CGG	CAG	GTC	GCC	GTC	GAC	GAG	1511	
Ala	Ser	Ala	Gly	Asp	Thr	Val	Thr	Phe	Arg	Gln	Val	Ala	Val	Asp	Glu		
			370						375				380				
GTC	GTT	GTG	ACC	AAC	AAC	CCC	GCC	GGC	GGC	GGC	AGC	GCC	CCC	ACC	TTC	1559	
Val	Val	Val	Thr	Asn	Asn	Pro	Ala	Gly	Gly	Gly	Ser	Ala	Pro	Thr	Phe		
			385					390					395				
ACC	GTG	AGA	GTG	CCC	CCT	TCA	AAC	GCT	TAC	ACC	AAC	ACC	GTG	TTT	AGG	1607	
Thr	Val	Arg	Val	Pro	Pro	Ser	Asn	Ala	Tyr	Thr	Asn	Thr	Val	Phe	Arg		
		400					405					410					
AAC	ACG	CTC	TTA	GAG	ACT	CGA	CCC	TCC	TCT	CGT	AGG	CTC	GAA	CTC	CCT	1655	
Asn	Thr	Leu	Leu	Glu	Thr	Arg	Pro	Ser	Ser	Arg	Arg	Leu	Glu	Leu	Pro		
		415				420					425				430		
ATG	CCA	CCT	GCT	GAC	TTT	GGA	CAG	ACG	GTC	GCC	AAC	AAC	CCG	AAG	ATC	1703	
Met	Pro	Pro	Ala	Asp	Phe	Gly	Gln	Thr	Val	Ala	Asn	Asn	Pro	Lys	Ile		
				435						440				445			
GAG	CAG	TCG	CTT	CTT	AAA	GAA	ACA	CTT	GGC	TGC	TAT	TTG	GTC	CAC	TCC	1751	

1

(2) INFORMATION FOR SEQ ID NO:50:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 647 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:50:

```

Met Gly Asp Ala Gly Val Ala Ser Gln Arg Pro His Asn Arg Arg Gly
 1             5             10             15

Thr Arg Asn Val Arg Val Ser Ala Asn Thr Val Thr Val Asn Gly Arg
          20             25             30

Arg Asn Gln Arg Arg Arg Thr Gly Arg Gln Val Ser Pro Pro Asp Asn
          35             40             45

Phe Thr Ala Ala Ala Gln Asp Leu Ala Gln Ser Leu Asp Ala Asn Thr
          50             55             60

Val Thr Phe Pro Ala Asn Ile Ser Ser Met Pro Glu Phe Arg Asn Trp
65             70             75             80

Ala Lys Gly Lys Ile Asp Leu Asp Ser Asp Ser Ile Gly Trp Tyr Phe
          85             90             95

Lys Tyr Leu Asp Pro Ala Gly Ala Thr Glu Ser Ala Arg Ala Val Gly
          100            105            110

Glu Tyr Ser Lys Ile Pro Asp Gly Leu Val Lys Phe Ser Val Asp Ala
          115            120            125

Glu Ile Arg Glu Ile Tyr Asn Glu Glu Cys Pro Val Val Thr Asp Val
          130            135            140

Ser Val Pro Leu Asp Gly Arg Gln Trp Ser Leu Ser Ile Phe Ser Phe
145            150            155            160

Pro Met Phe Arg Thr Ala Tyr Val Ala Val Ala Asn Val Glu Asn Lys
          165            170            175

Glu Met Ser Leu Asp Val Val Asn Asp Leu Ile Glu Trp Leu Asn Asn
          180            185            190

Leu Ala Asp Trp Arg Tyr Val Val Asp Ser Glu Gln Trp Ile Asn Phe
          195            200            205

Thr Asn Asp Thr Thr Tyr Tyr Val Arg Ile Arg Val Leu Arg Pro Thr
          210            215            220

Tyr Asp Val Pro Asp Pro Thr Glu Gly Leu Val Arg Thr Val Ser Asp
225            230            235            240

```

Tyr	Arg	Leu	Thr	Tyr	Lys	Ala	Ile	Thr	Cys	Glu	Ala	Asn	Met	Pro	Thr	245	250	255	
Leu	Val	Asp	Gln	Gly	Phe	Trp	Ile	Gly	Gly	Gln	Tyr	Ala	Leu	Thr	Pro	260	265	270	
Thr	Ser	Leu	Pro	Gln	Tyr	Asp	Val	Ser	Glu	Ala	Tyr	Ala	Leu	His	Thr	275	280	285	
Leu	Thr	Phe	Ala	Arg	Pro	Ser	Ser	Ala	Ala	Ala	Leu	Ala	Phe	Val	Trp	290	295	300	
Ala	Gly	Leu	Pro	Gln	Gly	Gly	Thr	Ala	Pro	Ala	Gly	Thr	Pro	Ala	Trp	305	310	315	320
Glu	Gln	Ala	Ser	Ser	Gly	Gly	Tyr	Leu	Thr	Trp	Arg	His	Asn	Gly	Thr	325	330	335	
Thr	Phe	Pro	Ala	Gly	Ser	Val	Ser	Tyr	Val	Leu	Pro	Glu	Gly	Phe	Ala	340	345	350	
Leu	Glu	Arg	Tyr	Asp	Pro	Asn	Asp	Gly	Ser	Trp	Thr	Asp	Phe	Ala	Ser	355	360	365	
Ala	Gly	Asp	Thr	Val	Thr	Phe	Arg	Gln	Val	Ala	Val	Asp	Glu	Val	Val	370	375	380	
Val	Thr	Asn	Asn	Pro	Ala	Gly	Gly	Gly	Ser	Ala	Pro	Thr	Phe	Thr	Val	385	390	395	400
Arg	Val	Pro	Pro	Ser	Asn	Ala	Tyr	Thr	Asn	Thr	Val	Phe	Arg	Asn	Thr	405	410	415	
Leu	Leu	Glu	Thr	Arg	Pro	Ser	Ser	Arg	Arg	Leu	Glu	Leu	Pro	Met	Pro	420	425	430	
Pro	Ala	Asp	Phe	Gly	Gln	Thr	Val	Ala	Asn	Asn	Pro	Lys	Ile	Glu	Gln	435	440	445	
Ser	Leu	Leu	Lys	Glu	Thr	Leu	Gly	Cys	Tyr	Leu	Val	His	Ser	Lys	Met	450	455	460	
Arg	Asn	Pro	Val	Phe	Gln	Leu	Thr	Pro	Ala	Ser	Ser	Phe	Gly	Ala	Val	465	470	475	480
Ser	Phe	Asn	Asn	Pro	Gly	Tyr	Glu	Arg	Thr	Arg	Asp	Leu	Pro	Asp	Tyr	485	490	495	
Thr	Gly	Ile	Arg	Asp	Ser	Phe	Asp	Gln	Asn	Met	Ser	Thr	Ala	Val	Ala	500	505	510	
His	Phe	Arg	Ser	Leu	Ser	His	Ser	Cys	Ser	Ile	Val	Thr	Lys	Thr	Tyr	515	520	525	
Gln	Gly	Trp	Glu	Gly	Val	Thr	Asn	Val	Asn	Thr	Pro	Phe	Gly	Gln	Phe	530	535	540	

Ala His Ala Gly Leu Leu Lys Asn Glu Glu Ile Leu Cys Leu Ala Asp
 545 550 555 560

Asp Leu Ala Thr Arg Leu Thr Gly Val Tyr Pro Ala Thr Asp Asn Phe
 565 570 575

Ala Ala Ala Val Ser Ala Phe Ala Ala Asn Met Leu Ser Ser Val Leu
 580 585 590

Lys Ser Glu Ala Thr Ser Ser Ile Ile Lys Ser Val Gly Glu Thr Ala
 595 600 605

Val Gly Ala Ala Gln Ser Gly Leu Ala Lys Leu Pro Gly Leu Leu Met
 610 615 620

Ser Val Pro Gly Lys Ile Ala Ala Arg Val Arg Ala Arg Arg Ala Arg
 625 630 635 640

Arg Arg Ala Ala Arg Ala Asn
 645

(2) INFORMATION FOR SEQ ID NO:51:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2479 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: unknown
- (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: DNA

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 283..2307

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:51:

GT TTTTCTTT CTTTACCAAG TGTGGTAAAA TTAAACAAA GAAGAAAACC AGGACCGTAA 60

CCCGGCCCTT ACACACCTCG AGTCCGTGAC CACCGGATTA TACGTCGCCC ACCACACGGC 120

GCCTTTTCCG ACCACTCTCG AGAGTCGTTG GGAGTTTCGT CCGTGACCAC CCGGTTGGCA 180

GTCGACAGAC GCTTCCGGAC CACTAGAACC TCCTCGAGCG ACGCACACAC AGCACACACA 240

CCGCCTTAGC TGCACCTACG GCAGCGTTGA TAGCGCGGAT TT ATG AGC GAG CAC 294
 Met Ser Glu His
 1

ACC ATC GCC CAC TCC ATC ACA TTA CCA CCC GGT TAC ACC CTT GCC CTA 342
 Thr Ile Ala His Ser Ile Thr Leu Pro Pro Gly Tyr Thr Leu Ala Leu
 5 10 15 20

ATA CCC CCT GAA CCT GAA GCA GGA TGG GAG ATG CTG GAG TGG CGT CAC 390
 Ile Pro Pro Glu Pro Glu Ala Gly Trp Glu Met Leu Glu Trp Arg His
 25 30 35

AGC	GAC	CTC	ACA	ACC	GTC	GCG	GAA	CCC	GTA	ACG	TTC	GGG	TCA	GCG	CCA	438
Ser	Asp	Leu	Thr	Thr	Val	Ala	Glu	Pro	Val	Thr	Phe	Gly	Ser	Ala	Pro	
			40					45					50			
ACA	CCG	TCA	CCG	TCA	ATG	GTA	GAA	GAA	ACC	AAC	GGC	GTC	GGA	CCG	GAA	486
Thr	Pro	Ser	Pro	Ser	Met	Val	Glu	Glu	Thr	Asn	Gly	Val	Gly	Pro	Glu	
		55					60					65				
GGC	AAG	TTT	CTC	CCC	CTG	ACA	ATT	TCA	CCG	CTG	CTG	CAC	AAG	ACC	TCG	534
Gly	Lys	Phe	Leu	Pro	Leu	Thr	Ile	Ser	Pro	Leu	Leu	His	Lys	Thr	Ser	
	70					75					80					
CGC	AAA	GCC	TTG	ACG	CCA	ACA	CCG	TCA	CTT	TCC	CCC	GCT	AAC	ATC	TCT	582
Arg	Lys	Ala	Leu	Thr	Pro	Thr	Pro	Ser	Leu	Ser	Pro	Ala	Asn	Ile	Ser	
85					90					95					100	
AGC	ATG	CCC	GAA	TTC	CGG	AAT	TGG	GCC	AAG	GGA	AAG	ATC	GAC	CTC	GAC	630
Ser	Met	Pro	Glu	Phe	Arg	Asn	Trp	Ala	Lys	Gly	Lys	Ile	Asp	Leu	Asp	
				105					110					115		
TCC	GAT	TCC	ATC	GGC	TGG	TAC	TTC	AAG	TAC	CTT	GAC	CCA	GCG	GGT	GCT	678
Ser	Asp	Ser	Ile	Gly	Trp	Tyr	Phe	Lys	Tyr	Leu	Asp	Pro	Ala	Gly	Ala	
			120					125					130			
ACA	GAG	TCT	GCG	CGC	GCC	GTC	GGC	GAG	TAC	TCG	AAG	ATC	CCT	GAC	GGC	726
Thr	Glu	Ser	Ala	Arg	Ala	Val	Gly	Glu	Tyr	Ser	Lys	Ile	Pro	Asp	Gly	
		135					140					145				
CTC	GTC	AAG	TTC	TCC	GTC	GAC	GCA	GAG	ATA	AGA	GAG	ATC	TAT	AAC	GAG	774
Leu	Val	Lys	Phe	Ser	Val	Asp	Ala	Glu	Ile	Arg	Glu	Ile	Tyr	Asn	Glu	
	150					155					160					
GAG	TGC	CCC	GTC	GTC	ACT	GAC	GTG	TCC	GTC	CCC	CTC	GAC	GGC	CGC	CAG	822
Glu	Cys	Pro	Val	Val	Thr	Asp	Val	Ser	Val	Pro	Leu	Asp	Gly	Arg	Gln	
165					170					175					180	
TGG	AGC	CTC	TCG	ATT	TTC	TCC	TTT	CCG	ATG	TTC	AGA	ACC	GCC	TAC	GTC	870
Trp	Ser	Leu	Ser	Ile	Phe	Ser	Phe	Pro	Met	Phe	Arg	Thr	Ala	Tyr	Val	
				185					190					195		
GCC	GTA	GCG	AAC	GTC	GAG	AAC	AAG	GAG	ATG	TCG	CTC	GAC	GTT	GTC	AAC	918
Ala	Val	Ala	Asn	Val	Glu	Asn	Lys	Glu	Met	Ser	Leu	Asp	Val	Val	Asn	
			200					205					210			
GAC	CTC	ATC	GAG	TGG	CTC	AAC	AAT	CTC	GCC	GAC	TGG	CGT	TAT	GTC	GTT	966
Asp	Leu	Ile	Glu	Trp	Leu	Asn	Asn	Leu	Ala	Asp	Trp	Arg	Tyr	Val	Val	
		215					220					225				
GAC	TCT	GAA	CAG	TGG	ATT	AAC	TTC	ACC	AAT	GAC	ACC	ACG	TAC	TAC	GTC	1014
Asp	Ser	Glu	Gln	Trp	Ile	Asn	Phe	Thr	Asn	Asp	Thr	Thr	Tyr	Tyr	Val	
	230					235					240					
CGC	ATC	CGC	GTT	CTA	CGT	CCA	ACC	TAC	GAC	GTT	CCA	GAC	CCC	ACA	GAG	1062
Arg	Ile	Arg	Val	Leu	Arg	Pro	Thr	Tyr	Asp	Val	Pro	Asp	Pro	Thr	Glu	
245					250					255					260	

GGC	CTT	GTT	CGC	ACA	GTC	TCA	GAC	TAC	CGC	CTC	ACT	TAT	AAG	GCG	ATA	1110
Gly	Leu	Val	Arg	Thr	Val	Ser	Asp	Tyr	Arg	Leu	Thr	Tyr	Lys	Ala	Ile	
				265					270					275		
ACA	TGT	GAA	GCC	AAC	ATG	CCA	ACA	CTC	GTC	GAC	CAA	GGC	TTT	TGG	ATC	1158
Thr	Cys	Glu	Ala	Asn	Met	Pro	Thr	Leu	Val	Asp	Gln	Gly	Phe	Trp	Ile	
			280					285					290			
GGC	GGC	CAG	TAC	GCT	CTC	ACC	CCG	ACT	AGC	CTA	CCG	CAG	TAC	GAC	GTC	1206
Gly	Gly	Gln	Tyr	Ala	Leu	Thr	Pro	Thr	Ser	Leu	Pro	Gln	Tyr	Asp	Val	
		295					300					305				
AGC	GAG	GCC	TAC	GCT	CTG	CAC	ACT	TTG	ACC	TTC	GCC	AGA	CCA	TCC	AGC	1254
Ser	Glu	Ala	Tyr	Ala	Leu	His	Thr	Leu	Thr	Phe	Ala	Arg	Pro	Ser	Ser	
	310					315					320					
GCC	GCT	GCA	CTC	GCG	TTT	GTG	TGG	GCA	GGT	TTG	CCA	CAG	GGT	GGC	ACT	1302
Ala	Ala	Ala	Leu	Ala	Phe	Val	Trp	Ala	Gly	Leu	Pro	Gln	Gly	Gly	Thr	
325				330					335						340	
GCG	CCT	GCA	GGC	ACT	CCA	GCC	TGG	GAG	CAG	GCA	TCC	TCG	GGT	GGC	TAC	1350
Ala	Pro	Ala	Gly	Thr	Pro	Ala	Trp	Glu	Gln	Ala	Ser	Ser	Gly	Gly	Tyr	
				345				350					355			
CTC	ACC	TGG	CGC	CAC	AAC	GGT	ACT	ACT	TTC	CCA	GCT	GGC	TCC	GTT	AGC	1398
Leu	Thr	Trp	Arg	His	Asn	Gly	Thr	Thr	Phe	Pro	Ala	Gly	Ser	Val	Ser	
			360				365						370			
TAC	GTT	CTC	CCT	GAG	GGT	TTC	GCC	CTT	GAG	CGC	TAC	GAC	CCG	AAC	GAC	1446
Tyr	Val	Leu	Pro	Glu	Gly	Phe	Ala	Leu	Glu	Arg	Tyr	Asp	Pro	Asn	Asp	
		375				380						385				
GGC	TCT	TGG	ACC	GAC	TTC	GCT	TCC	GCA	GGA	GAC	ACC	GTC	ACT	TTC	CGG	1494
Gly	Ser	Trp	Thr	Asp	Phe	Ala	Ser	Ala	Gly	Asp	Thr	Val	Thr	Phe	Arg	
	390				395						400					
CAG	GTC	GCC	GTC	GAC	GAG	GTC	GTT	GTG	ACC	AAC	AAC	CCC	GCC	GGC	GGC	1542
Gln	Val	Ala	Val	Asp	Glu	Val	Val	Val	Thr	Asn	Asn	Pro	Ala	Gly	Gly	
405				410					415					420		
GGC	AGC	GCC	CCC	ACC	TTC	ACC	GTG	AGA	GTG	CCC	CCT	TCA	AAC	GCT	TAC	1590
Gly	Ser	Ala	Pro	Thr	Phe	Thr	Val	Arg	Val	Pro	Pro	Ser	Asn	Ala	Tyr	
			425				430					435				
ACC	AAC	ACC	GTG	TTT	AGG	AAC	ACG	CTC	TTA	GAG	ACT	CGA	CCC	TCC	TCT	1638
Thr	Asn	Thr	Val	Phe	Arg	Asn	Thr	Leu	Leu	Glu	Thr	Arg	Pro	Ser	Ser	
			440				445					450				
CGT	AGG	CTC	GAA	CTC	CCT	ATG	CCA	CCT	GCT	GAC	TTT	GGA	CAG	ACG	GTC	1686
Arg	Arg	Leu	Glu	Leu	Pro	Met	Pro	Pro	Ala	Asp	Phe	Gly	Gln	Thr	Val	
		455				460						465				
GCC	AAC	AAC	CCG	AAG	ATC	GAG	CAG	TCG	CTT	CTT	AAA	GAA	ACA	CTT	GGC	1734
Ala	Asn	Asn	Pro	Lys	Ile	Glu	Gln	Ser	Leu	Leu	Lys	Glu	Thr	Leu	Gly	
	470				475						480					

TGC TAT TTG GTC CAC TCC AAA ATG CGA AAC CCC GTT TTC CAG CTC ACG Cys Tyr Leu Val His Ser Lys Met Arg Asn Pro Val Phe Gln Leu Thr 485 490 495 500	1782
CCA GCC AGC TCC TTT GGC GCC GTT TCC TTC AAC AAT CCG GGT TAT GAG Pro Ala Ser Ser Phe Gly Ala Val Ser Phe Asn Asn Pro Gly Tyr Glu 505 510 515	1830
CGC ACA CGC GAC CTC CCG GAC TAC ACT GGC ATC CGT GAC TCA TTC GAC Arg Thr Arg Asp Leu Pro Asp Tyr Thr Gly Ile Arg Asp Ser Phe Asp 520 525 530	1878
CAG AAC ATG TCC ACC GCT GTG GCC CAC TTC CGC TCA CTC TCC CAC TCC Gln Asn Met Ser Thr Ala Val Ala His Phe Arg Ser Leu Ser His Ser 535 540 545	1926
TGC AGT ATC GTC ACT AAG ACC TAC CAG GGT TGG GAA GGC GTC ACG AAC Cys Ser Ile Val Thr Lys Thr Tyr Gln Gly Trp Glu Gly Val Thr Asn 550 555 560	1974
GTC AAC ACG CCT TTC GGC CAA TTC GCG CAC GCG GGC CTC CTC AAG AAT Val Asn Thr Pro Phe Gly Gln Phe Ala His Ala Gly Leu Leu Lys Asn 565 570 575 580	2022
GAG GAG ATC CTC TGC CTC GCC GAC GAC CTG GCC ACC CGT CTC ACA GGT Glu Glu Ile Leu Cys Leu Ala Asp Asp Leu Ala Thr Arg Leu Thr Gly 585 590 595	2070
GTC TAC CCC GCC ACT GAC AAC TTC GCG GCC GCC GTT TCT GCC TTC GCC Val Tyr Pro Ala Thr Asp Asn Phe Ala Ala Val Ser Ala Phe Ala 600 605 610	2118
GCG AAC ATG CTG TCC TCC GTG CTG AAG TCG GAG GCA ACG TCC TCC ATC Ala Asn Met Leu Ser Ser Val Leu Lys Ser Glu Ala Thr Ser Ser Ile 615 620 625	2166
ATC AAG TCC GTT GGC GAG ACT GCC GTC GGC GCG GCT CAG TCC GGC CTC Ile Lys Ser Val Gly Glu Thr Ala Val Gly Ala Ala Gln Ser Gly Leu 630 635 640	2214
GCG AAG CTA CCC GGA CTG CTA ATG AGT GTA CCA GGG AAG ATT GCC GCG Ala Lys Leu Pro Gly Leu Leu Met Ser Val Pro Gly Lys Ile Ala Ala 645 650 655 660	2262
CGT GTC CGC GCG CGC CGA GCG CGC CGC CGC GCT CGT GCC AAT Arg Val Arg Ala Arg Arg Ala Arg Arg Ala Ala Arg Ala Asn 665 670 675	2307
TAGTTTGCTC GCTCCTGTTT CGCCGTTTCG TAAAACGGCG TGGTCCCGCA CATTACGCGT	2367
ACCCCTAAAGA CTCTGGTGAG TCCCCGTCGT TACACGACGG GTCTGCCGCG GTTCGATTCC	2427
ATTCCCAAGC GGCAAGAAGG ACGTAGTTAG CTCTGCGTCC CTCGGGATAC CA	2479

(2) INFORMATION FOR SEQ ID NO:52:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 675 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:52:

```

Met Ser Glu His Thr Ile Ala His Ser Ile Thr Leu Pro Pro Gly Tyr
 1             5             10             15

Thr Leu Ala Leu Ile Pro Pro Glu Pro Glu Ala Gly Trp Glu Met Leu
      20             25             30

Glu Trp Arg His Ser Asp Leu Thr Thr Val Ala Glu Pro Val Thr Phe
      35             40             45

Gly Ser Ala Pro Thr Pro Ser Pro Ser Met Val Glu Glu Thr Asn Gly
      50             55             60

Val Gly Pro Glu Gly Lys Phe Leu Pro Leu Thr Ile Ser Pro Leu Leu
      65             70             75             80

His Lys Thr Ser Arg Lys Ala Leu Thr Pro Thr Pro Ser Leu Ser Pro
      85             90             95

Ala Asn Ile Ser Ser Met Pro Glu Phe Arg Asn Trp Ala Lys Gly Lys
      100            105            110

Ile Asp Leu Asp Ser Asp Ser Ile Gly Trp Tyr Phe Lys Tyr Leu Asp
      115            120            125

Pro Ala Gly Ala Thr Glu Ser Ala Arg Ala Val Gly Glu Tyr Ser Lys
      130            135            140

Ile Pro Asp Gly Leu Val Lys Phe Ser Val Asp Ala Glu Ile Arg Glu
      145            150            155            160

Ile Tyr Asn Glu Glu Cys Pro Val Val Thr Asp Val Ser Val Pro Leu
      165            170            175

Asp Gly Arg Gln Trp Ser Leu Ser Ile Phe Ser Phe Pro Met Phe Arg
      180            185            190

Thr Ala Tyr Val Ala Val Ala Asn Val Glu Asn Lys Glu Met Ser Leu
      195            200            205

Asp Val Val Asn Asp Leu Ile Glu Trp Leu Asn Asn Leu Ala Asp Trp
      210            215            220

Arg Tyr Val Val Asp Ser Glu Gln Trp Ile Asn Phe Thr Asn Asp Thr
      225            230            235            240

```

Thr Tyr Tyr Val Arg Ile Arg Val Leu Arg Pro Thr Tyr Asp Val Pro
 245 250 255
 Asp Pro Thr Glu Gly Leu Val Arg Thr Val Ser Asp Tyr Arg Leu Thr
 260 265 270
 Tyr Lys Ala Ile Thr Cys Glu Ala Asn Met Pro Thr Leu Val Asp Gln
 275 280 285
 Gly Phe Trp Ile Gly Gly Gln Tyr Ala Leu Thr Pro Thr Ser Leu Pro
 290 295 300
 Gln Tyr Asp Val Ser Glu Ala Tyr Ala Leu His Thr Leu Thr Phe Ala
 305 310 315 320
 Arg Pro Ser Ser Ala Ala Ala Leu Ala Phe Val Trp Ala Gly Leu Pro
 325 330 335
 Gln Gly Gly Thr Ala Pro Ala Gly Thr Pro Ala Trp Glu Gln Ala Ser
 340 345 350
 Ser Gly Gly Tyr Leu Thr Trp Arg His Asn Gly Thr Thr Phe Pro Ala
 355 360 365
 Gly Ser Val Ser Tyr Val Leu Pro Glu Gly Phe Ala Leu Glu Arg Tyr
 370 375 380
 Asp Pro Asn Asp Gly Ser Trp Thr Asp Phe Ala Ser Ala Gly Asp Thr
 385 390 395 400
 Val Thr Phe Arg Gln Val Ala Val Asp Glu Val Val Val Thr Asn Asn
 405 410 415
 Pro Ala Gly Gly Gly Ser Ala Pro Thr Phe Thr Val Arg Val Pro Pro
 420 425 430
 Ser Asn Ala Tyr Thr Asn Thr Val Phe Arg Asn Thr Leu Leu Glu Thr
 435 440 445
 Arg Pro Ser Ser Arg Arg Leu Glu Leu Pro Met Pro Pro Ala Asp Phe
 450 455 460
 Gly Gln Thr Val Ala Asn Asn Pro Lys Ile Glu Gln Ser Leu Leu Lys
 465 470 475 480
 Glu Thr Leu Gly Cys Tyr Leu Val His Ser Lys Met Arg Asn Pro Val
 485 490 495
 Phe Gln Leu Thr Pro Ala Ser Ser Phe Gly Ala Val Ser Phe Asn Asn
 500 505 510
 Pro Gly Tyr Glu Arg Thr Arg Asp Leu Pro Asp Tyr Thr Gly Ile Arg
 515 520 525
 Asp Ser Phe Asp Gln Asn Met Ser Thr Ala Val Ala His Phe Arg Ser
 530 535 540

Leu Ser His Ser Cys Ser Ile Val Thr Lys Thr Tyr Gln Gly Trp Glu
545 550 555 560

Gly Val Thr Asn Val Asn Thr Pro Phe Gly Gln Phe Ala His Ala Gly
565 570 575

Leu Leu Lys Asn Glu Glu Ile Leu Cys Leu Ala Asp Asp Leu Ala Thr
580 585 590

Arg Leu Thr Gly Val Tyr Pro Ala Thr Asp Asn Phe Ala Ala Ala Val
595 600 605

Ser Ala Phe Ala Ala Asn Met Leu Ser Ser Val Leu Lys Ser Glu Ala
610 615 620

Thr Ser Ser Ile Ile Lys Ser Val Gly Glu Thr Ala Val Gly Ala Ala
625 630 635 640

Gln Ser Gly Leu Ala Lys Leu Pro Gly Leu Leu Met Ser Val Pro Gly
645 650 655

Lys Ile Ala Ala Arg Val Arg Ala Arg Arg Ala Arg Arg Arg Ala Ala
660 665 670

Arg Ala Asn
675

(2) INFORMATION FOR SEQ ID NO:53:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 59 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: unknown
 - (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:53:

GGGGATCCAC AGTTCTGCCT CCCCCGGACG GTAAATATAG GGGAACCATG GTCTAGAGG

59

(2) INFORMATION FOR SEQ ID NO:54:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 29 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: unknown
 - (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:54:

GGCCGCTTAA TTAAGGATCC GCGCGCCA

29

(2) INFORMATION FOR SEQ ID NO:55:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 29 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: unknown
 - (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:55:

CGAATTAATT CCTAGGCCGC GCGGTGATC

29

(2) INFORMATION FOR SEQ ID NO:56:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 8 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: unknown
 - (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:56:

TTAATTAA

8

(2) INFORMATION FOR SEQ ID NO:57:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 8 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: unknown
 - (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:57:

GGCGCGCC

8